


```
Db 1 MAAQPLRHSRCATPRGDFCGGTERAIDQASFTTSMEWDQVVGKSSPLGAGLGAERP 60
QY 61 AAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVPSRVTVNNVLEAPFL 120
Db 61 AAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVPSRVTVNNVLEAPFL 120
QY 121 VGTEGLKSTYNLPCGSGGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLTKTAIV 180
Db 121 VGTEGLKSTYNLPCGSGGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLTKTAIV 180
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 5
US-09-942-052-707
; Sequence 707, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Splice variant 1
; OTHER INFORMATION: open reading frame 3 peptide sequence
; NAME/KEY: MOD RES
; LOCATION: (44)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (50)..(51)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (67)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (93)
; OTHER INFORMATION: Unknown amino acid or stop
US-09-942-052-707

Query Match 11.6%; Score 138.5; DB 10; Length 164;
Best Local Similarity 82.1%; Pred. No. 9e-06;
Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 162 FCLSSDKMVCYLLTKTAIVNASEMDIQNVPLSEKIAELK 200
Db 2 FFLSS-----YLLTKTAIVNASEMDIQNVPLSEKIAELK 35

RESULT 6
US-09-942-052-704
; Sequence 704, Application US/09942052
; Publication No. US20030170626A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 704
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3 peptide
; OTHER INFORMATION: sequence
US-09-942-052-704

Query Match 11.5%; Score 138; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 YLLTKTAIVNASEMDIQNVPLSEKIAELK 200
Db 1 YLLTKTAIVNASEMDIQNVPLSEKIAELK 29

RESULT 7
US-09-890-688-82
; Sequence 82, Application US/09890688
; Publication No. US2003014475A1
; GENERAL INFORMATION:
; APPLICANT: Seishi KATO
; APPLICANT: Chikashi EGUCHI
; APPLICANT: Mihoro SAEKI
; TITLE OF INVENTION: Human Proteins and cDNAs thereof
; FILE REFERENCE: 2001-1102A/WMC/00653
; CURRENT APPLICATION NUMBER: US/09/890,688
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: JP 11-346863
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 11-34684
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 2000-31062
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: JP 2000-34091
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-34090
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-35829
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-35899
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-71161
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: JP 2000-160851
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-890-688-82

Query Match 9.4%; Score 113; DB 10; Length 233;
```



```
Db      89 EDSPQPGVPLILLP-----LCRCYAKEICSEYVVRTTDLVNH-----LNSNAIS 135
QY      116 EAPFLVIGIEGSLKSTYNNLLFCGSCGIPVGFHLYSTHAALALRGHFCLLSSDKMVCYLK 175
Db      136 EGGP-----SMRKARFLGSAS---AFSVKQTEWP-----CATTSKKCY-IQ 173
QY      176 TKAIVNASEMDIQN-----VPLSEKIAELKEKIVLTH 207
Db      174 NGSFGGITDEEQSSLYNLYPSKELLPPDKEMSI FDH 211

RESULT 14
US-10-437-963-155606
; Sequence 155606, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155606
; LENGTH: 848
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(848)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55353C.1.pap
US-10-437-963-155606

Query Match      7.3%; Score 87.5; DB 16; Length 848;
Best Local Similarity 22.1%; Pred. No. 14;
Matches 65; Conservative 23; Mismatches 109; Indels 97; Gaps 13;

QY      16 PRGDFCGGTERAIDQASFTTSMEWDTQVVKGSPGLGAEAP-----60
Db      85 PRGFRLDGTGCGVRQTASCHLTDLTDLKSGCLGPRGRKRNPTTPLLGGNQBEGRHL 144
QY      61 AAGPOLPSMLOPERCAVFCQACH-----AVLADSVHLAWDLGR-----99
Db      145 TLGPDVPEGARPRGLCLPQAGSHDTPGPISPTSLVLRKTRIEDVAVTMDISEANEG 204
QY      100 --SLGAVV-FSRVTNNV-----VLEAPFLVIGIEGSLK-GSTYNLLFCG-----138
Db      205 YVSCGSVIEWSQMKAAARAGVRAQSPACKXGIPTVGVKGSV--LLFTARMGEGLCHEMTT 262
QY      139 -----SCGIPV-----GFHLYSTH---AALALRGHFCLLSSDKMVCYL 173
Db      263 GSKAMECGVKMCLVGNPWLAWDELGRPSGYQFGLNHRRLPVKAVLDGHLCEEDAVNCLP 322
QY      174 LKTKAIVNASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQSKP 227
Db      323 KLRKTVGSASE-----AKEAVKPAVKQEKIRS-VKVLVLVS-DLSLP 363
```

RESULT 15
US-10-282-122A-51802
; Sequence 51802, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

```
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51802
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51802

Query Match      7.2%; Score 86.5; DB 15; Length 669;
Best Local Similarity 27.4%; Pred. No. 13;
Matches 37; Conservative 22; Mismatches 45; Indels 31; Gaps 8;

QY      96 DLGR---SIGAVVFSRVTNNVLEAPFLVIGIEGSLKSTYNL-----LFCGSCGIPVGF 146
Db      366 DIGRKKVKIGSRVFRSNDVI---PEIMGVTEETEGETNEIEAETICPCGSEIVKEGV 422
QY      147 HLYSTHAALALRGHFCLLSSDKMVCYLKTKAIVNASEMDIQNV-LSEKIAE-LKEKIV 204
Db      423 HL-----FC--ENTLSCKPQWKSIVHFASREAMNIEGFSEKTAQLFEK---465
QY      205 LTHNELKSLMKILSE 219
Db      466 LNIKSIDLYRITKE 480

Search completed: May 27, 2005, 09:05:36
Job time : 138 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 08:43:04 ; Search time 40 Seconds
(without alignments)
550.841 Million cell updates/sec

Title: US-09-942-052A-728
Perfect score: 1198
Sequence: 1 MAAQPLRHRSRCATPRGRDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	7.7	695	2 C86731	copper-potassium t
2	90.5	7.6	597	2 D71293	probable phosphori
3	89	7.4	306	2 T46399	hypothetical prote
4	88	7.3	250	2 A43623	kanamycin kinase (
5	86.5	7.2	669	2 A97229	NAD-dependent DNA
6	85	7.1	2194	1 JQ1877	glutamate synthase
7	84.5	7.1	361	2 AE2074	ferrichrome bindin
8	83	6.9	663	2 D97047	DNA ligase (NAD de
9	82.5	6.9	143	2 B69099	formate hydrogenly
10	82.5	6.9	493	2 A85433	sugar transporter
11	82	6.8	967	2 F87678	DNA polymerase I (
12	81.5	6.8	329	2 D90404	transport protein,
13	81.5	6.8	457	2 H85767	glucuronide permea
14	81.5	6.8	457	2 C90919	glucuronide permea
15	81.5	6.8	457	2 B64918	glucuronide permea
16	81.5	6.8	640	2 T41977	hypothetical prote
17	81.5	6.8	662	2 T44221	probable capsid'pr
18	81	6.8	1418	2 S40764	hypothetical prote
19	80.5	6.7	766	2 B85440	receptor kinase-li
20	80.5	6.7	1711	1 A55148	protein-tyrosine-p
21	80	6.7	469	2 S55167	IME2-dependent sig
22	79	6.6	386	2 H90789	probable aminometh
23	79	6.6	386	2 C85650	probable aminometh
24	79	6.6	437	2 C86823	GTP-binding protei
25	78.5	6.6	188	2 F83816	late competence op
26	78.5	6.6	578	2 F86484	probable hydroxyme
27	78.5	6.6	697	1 QQBK3	UL104 protein - hu
28	78	6.5	351	2 H82098	UDP-3-O-3-hydroxym
29	77.5	6.5	662	2 T44036	hypothetical prote

30	77.5	6.5	1013	2 T31211	trwC protein homol
31	77.5	6.5	1643	2 T14274	versican precursor
32	77	6.4	381	2 AD113	hypothetical prote
33	77	6.4	721	2 F87611	TonB-dependent rec
34	77	6.4	738	1 TFHUM	melanotransferrin
35	76.5	6.4	307	2 AG2017	glycerol-3-phospha
36	76.5	6.4	387	2 D69392	probable acyl-CoA
37	76.5	6.4	388	2 S57526	cellulase - Fibrob
38	76.5	6.4	764	2 AD3144	formate dehydrogen
39	76.5	6.4	764	2 H98143	cbbc protein (U60
40	76.5	6.4	1554	2 T06370	probable DNA (cyto
41	76	6.3	367	2 AH0936	glycerol dehydroge
42	75.5	6.3	310	2 A55053	endothelial monocy
43	75.5	6.3	415	2 T48716	hypothetical prote
44	75.5	6.3	700	2 D70951	probable UvrD - My
45	75.5	6.3	1534	2 S59604	DNA (cytosine-5-) -

ALIGNMENTS

RESULT 1

C86731
copper-potassium transporting ATPase B copB [imported] - Lactococcus lactis subsp. lactis
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86731
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A:Reference number: A86625; UID:21235186; PMID:11337471
A:Accession: C86731
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-695 <STO>
A:Cross-references: UNIPROT:Q5CH87; GB:AE005176; PID:g12723778; PIDN:AAK04949.1; GSPDB:GN
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: copB
C:Superfamily: Enterococcus copper-transporting ATPase copB; ATPase nucleotide-binding do

Query Match	7.7%	Score 92;	DB 2;	Length 695;
Best Local Similarity	25.3%	Pred. No. 2;		
Matches	38;	Conservative	28;	Mismatches 50; Indels 34; Gaps 7;
QY	101	LGAVFVSRTNNVLEAPFLVGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRG	160	
Db	101	LGTTIIF-----FYSTGTPFFSGAKGELSKRPANMMLITWGITVAY-AYSVYATIMSLNG	153	
QY	161	HP-----CLSSDKMVCYLLKTKAIVNASE--MDI-----QNVPLSE-K	195	
Db	154	HGMNFWFELATLIVIMLIGHLIEMKAIMGAGDALKDLSLVPKKAHLKSGKDVELSELK	213	
QY	196	IABLKEKIVLTNRKLSMKILSEVTPDOS	225	
Db	214	VGDL--LLVKEKEKIPADGLILSEALVDES	241	

RESULT 2

D71293
probable phosphoribosylglycinamide formyltransferase - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: D71293
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
rson, J.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; UID:98332770; PMID:9665876
A:Accession: D71293
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A;Residues: 1-597 <COL>
A;Cross-references: UNIPROT:O83693; GB:AE001243; GB:AE000520; NID:g3322990; PIDN:AA6566
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0695

Query Match 7.6%; Score 90.5; DB 2; Length 597;
Best Local Similarity 28.0%; Pred. No. 2.3;
Matches 51; Conservative 19; Mismatches 47; Indels 65; Gaps 12;

QY 12 CATPPGDCGGTERAID-----QAST-----TSMEDTQVVGSSPLGPA 53
DB 90 CALP--GHRLEATKNAATKTRMRACFTRLARLCPRTFLEPDSFAWD-----PPGHA 140
QY 54 GLCAEPAAGPQLPSWLQP--ERCAVQC--AOCHAVLADSVHLAWLDSRLGAVVFSRVY 110
DB 141 RLCSHLHSAGLFFLVVKTDNMGARGCTLAQCKOTLINACAVARQFSRS-----190
QY 111 NNVLBAFLVIGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFLCLSDKMV 170
DB 191 GRVIE-EIFVGRFSLG-----LIFDGT-----LVVT--ALA-----DRHI 225
QY 171 CY 172
DB 226 CF 227

RESULT 3
T46399
hypothetical protein DKFp434N2420.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46399
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46399
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-306 <AAA>
A;Cross-references: UNIPROT:Q96EP1; EMBL:AL137561
A;Experimental source: adult testis; clone DKFp434N2420
C;Genetics:
A;Note: DKFp434N2420.1

Query Match 7.4%; Score 89; DB 2; Length 306;
Best Local Similarity 23.0%; Pred. No. 1.4;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPPRGDFC-----GTERAIDQA-----SFTTSMEDTQVVGSSPL-----50
DB 76 YRQMAQPPH---CPAPEGEPGAPQALGDAPSTSVSLTTAVQDYVCPLOGSHALCTCCFQ 132
QY 51 ----GPAGLGAEBPAGPQLPSWLQPERCAVQCAQCHAVLADSVHLAWLDSR-----99
DB 133 PMPDRAERQDPRAVQ-----QCAVC---LQPFCHLYMGCTRTGCVGCLIA 176
QY 100 ----SLGAVFSRVNTNNVLEAPFLVIGIEGSLKSTY--NLLFCGSCGIPVGFHLYSTHA 153
DB 177 PCLENLGDKLDGLVNNNSYEDILKNVLYAT--RGLTWKML-----TES 220
QY 154 ALAALRGHFLS-----SDRMVYLLKTKAIYNASEMDIQNVPLSE 194
DB 221 LVALQRGVLLSDYRVGTGTVLCYCCGLRSFRELTYVYRQNPASE 266

RESULT 4
A43623
kanamycin kinase (EC 2.7.1.95) - Campylobacter jejuni
C;Species: Campylobacter jejuni
C;Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 09-Jul-2004
C;Accession: A43623
R;Tenover, F.C.; Gilbert, T.; O'Hara, P.

Plasmid 22, 52-58, 1989
A;Title: Nucleotide sequence of a novel kanamycin resistance gene, aphA-7, from Campylobacter
A;Reference number: A43623; MUID:89387451; PMID:2550983
A;Accession: A43623
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-250 <TEN>
A;Cross-references: UNIPROT:P14508; GB:M29953; GB:J03316; NID:g144186; PIDN:AAA76822.1; I
C;Superfamily: kanamycin kinase
C;Keywords: phosphotransferase

Query Match 7.3%; Score 88; DB 2; Length 250;
Best Local Similarity 28.7%; Pred. No. 1.3;
Matches 33; Conservative 18; Mismatches 46; Indels 18; Gaps 5;

QY 105 VFSRVNTNNVLEAPFLVIGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRG-HFC 163
DB 40 IFSKTTYSVKREAWMMMLSDKLKVPDV-----IYGVRESEYILIMSELRGKHD 90
QY 164 LSSDKMVCYLLKTKAIYNA---SEMDIQNVPLSEKI-AELKEKIVLTHNRLKSL 213
DB 91 CFIDHPIKYI---ECLVNALHQLQDAIDRNCFPSSKIDVRLKELKYLDDNRIADI 142

RESULT 5
A97229
NAD-dependent DNA ligase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 12-Jul-2004
C;Accession: A97229
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A97229
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-669 <KUR>
A;Cross-references: UNIPROT:Q97FQ5; GB:AE001437; PIDN:AAK80620.1; PID:g15025704; GSPDB:GI
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2673
C;Superfamily: NAD+-dependent DNA ligase, Liga type

Query Match 7.2%; Score 86.5; DB 2; Length 669;
Best Local Similarity 27.4%; Pred. No. 6.4;
Matches 37; Conservative 22; Mismatches 45; Indels 31; Gaps 8;

QY 96 DLRS---SLGAVFSRVNTNNVLEAPFLVIGIEGSLKSTYNNL-----LFCGSCGIPVGF 146
DB 366 DIGRKKVIGSRVFRSNDVI---PEIMGVTEETEGETNEIETAPICPCGSEIVKEGV 422
QY 147 HLYSTHAALALRGHFLCSSDKMVCYLLKTKAIYNASEMDIQNV-LSEKIAE-LKEKIV 204
DB 423 HL-----FC--ENTLSCKPQWKSIVHFASREAMNIEGFSEKTAQLFEK--465
QY 205 LTHNRLKSLKILSE 219
DB 466 LNIKSIDLYRITKE 480

RESULT 6
JQ1977
glutamate synthase (NADH2) (EC 1.4.1.14) precursor [validated] - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Aug-2004
C;Accession: JQ1977; PQ0551
R;Gregerson, R.G.; Miller, S.S.; Twary, S.N.; Gantt, J.S.; Vance, C.P.
Plant Cell 5, 215-226, 1993
A;Title: Molecular characterization of NADH-dependent glutamate synthase from alfalfa noc
A;Reference number: JQ1977; MUID:93200806; PMID:8453303
A;Accession: JQ1977

A;Molecule type: mRNA
A;Residues: 1-2194 <GRE>
A;Cross-references: UNIPROT:Q03460; GB:L01660; NID:g166411; PIDN:AAB46617.1; PID:g166412
A;Accession: PQ0551
A;Molecule type: protein
A;Residues: 102-114 <GR2>
C;Comment: This enzyme catalyzes the reductive transfer of the amido group of glutamine
C;Superfamily: Glutamate synthase (NADH/NADPH), eukaryotic type
C;Keywords: 3Fe-4S; chloroplast; iron-sulfur protein; metalloprotein; NAD; oxidoreductase
F;1-101/Domain: propeptide #status predicted <PRO>
F;102-2194/Product: glutamate synthase (NADH) #status experimental <MAT>
F;102/Active site: Cys #status predicted
F;1246,1252,1257/Binding site: 3Fe-4S cluster (Cys) (covalent) #status predicted
Query Match 7.1%; Score 85; DB 1; Length 2194;
Best Local Similarity 24.7%; Pred. No. 40;
Matches 58; Conservative 22; Mismatches 79; Indels 76; Gaps 12;
QY 4 QPLRHRSRCATPRGDFCGGTERAIDQAS-----AAGPOLPSMLQPERCAVFOCAQCHAVLADSVHLAWDLRSR 99
Db 1037 EPLADGSR--NPKRS-----AIQVASGRFGVSSYYLTNADELQIKMAQAKP---- 1082
QY 54 GLCAEPP-----AAGPOLPSMLQPERCAVFOCAQCHAVLADSVHLAWDLRSR 99
Db 1083 GEGELPGHKVIGDIALTRNSTAGVGLIS--PPPHDIYS-----IEDLAQLIHLKXN 1133
QY 100 SLGAVFSRVNNVLEAPFLVIGIEGSLKSTYNLLFCG-----SCGIPVG 145
Db 1134 ANPA--ARISVKLVSEAGVGTASGVKGVKASHVLSIGHDGGTGASRWGTIGKSLPWE 1190
QY 146 FHLVSTHALLA--LQGHFCLSSDKWVCYLLKT-----KAIYNASEMDIQNVPL 192
Db 1191 LGLAETHQTLVANDLRGRTTLQDQ----LKTGRDVAIALIGABEYGFSTAPL 1241
RESULT 7
AE2074
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE2074
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-361 <KUR>
A;Cross-references: UNIPROT:Q8YV34; GB:BA000019; PIDN:BA073846.1; PID:g17131238; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2147
C;Superfamily: ferrichrome-iron transport protein fecB
Query Match 7.1%; Score 84.5; DB 2; Length 361;
Best Local Similarity 26.1%; Pred. No. 4.5;
Matches 36; Conservative 17; Mismatches 36; Indels 49; Gaps 7;
QY 97 LRSLSGAVP-----SRVT-----NNVLEAPFLVIGIEGSLKSTYNLLFCGSC-----GIPV 144
Db 81 ISHALGVKVIPLKQPVVLEENIILDSVLALGVK-----PVGVMYQDCENFRGIP- 133
QY 145 GFHLVSTHALLAALRGHFCCLSSDKWVCYLLKTKAIYNASEMDIQNVPLSKIAELAEKIV 204
Db 134 -----SD-----LLADVVG-----NIGNQPSLEKILSLKPDLI 163
QY 205 LTHNLKSLMKILSEVTP 222
Db 164 LGITWLKSSYKILSSIAP 181

RESULT 8

D97047
DNA ligase (NAD dependent), Liga [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 12-Jul-2004
C;Accession: D97047
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97047
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-663 <KUR>
A;Cross-references: UNIPROT:Q97JS8; GB:AE001437; PIDN:AAK79167.1; PID:g15024117; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1195
C;Superfamily: NAD+-dependent DNA ligase, Liga type

Query Match 6.9%; Score 83; DB 2; Length 663;
Best Local Similarity 28.6%; Pred. No. 13;
Matches 36; Conservative 18; Mismatches 44; Indels 28; Gaps 7;

QY 102 GAVFSRVNNVLEAPFLVIGIEGSLKSTYNL-----LFCGSCGIPVGFHLYSTHAAL 155

Db 368 GARVFLRRNDVI---PEIMGVTEETEGETEAPTCPCGSEIVKGVHL----- 417

QY 156 AALRGHFCCLSSDKWVCYLLKTKAIYNASEMDIQNVPL-LSEKIAE-LKEKIVLTHNLKSL 213

Db 418 -----FC--ENTLSCKPQVKSIVHFASRKAMNIEGFSEKTAQLFEK--LNIKSIDL 467

QY 214 MKILSE 219

Db 468 YRITKE 473

RESULT 9

B69099
C;Species: Methanobacterium thermoautotrophicum
C;Date: 03-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: B69099
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; I
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
A;Reference number: A69000; MUID:98037514; PMID:93711463
A;Accession: B69099
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-143 <MTH>
A;Cross-references: UNIPROT:Q27769; GB:AE0000929; GB:AE000666; NID:g2622853; PIDN:AAB8620
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1736
C;Superfamily: nrfC protein; ferredoxin 2[4Fe-4S] homology
F;82-118/Domain: ferredoxin 2[4Fe-4S] homology <FER2>

Query Match 6.9%; Score 82.5; DB 2; Length 143;
Best Local Similarity 23.0%; Pred. No. 2.1;
Matches 35; Conservative 26; Mismatches 66; Indels 25; Gaps 6;

QY 72 PERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVF-----SRVTNNVLEAPFLVIGIEG 125

Db 8 PELCD--ECMKERICPKNAIRVID-----GVPVFCMCSPERAPCLNICPEDAIVEVDG 60

QY 126 SLKSGSTYNLLFCGSC--GIPVGFHLYSTHAALALRGHFCCLSSDKWVCYLLKTKAIYNAS 183

Db 61 AVVILEDRICGICGLCRDACEVVG--AITLNERGVAVKCDLICDRDKPLCVWVCPKGLSLES 118


```
Query Match      6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 12;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVVFSRVTVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIASIGCGVTMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374

QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 375 QAIG---GSIPAFILGSLGVIANQVQTPVEVIMGIRTSIALVPCGFMLLAFVFIWFYPLTD 431

QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221
Db 432 K--KPKSEIVVEIDNRKKVQQQLISDIT 456

RESULT 14
C90919
glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C90919
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C90919
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-457 <HAY>
A;Cross-references: UNIPROT:Q8X673; GB:BA000007; PIDN:BA035746.1; PID:g13361790; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2323
C;Superfamily: melibiose carrier protein

Query Match      6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 12;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVVFSRVTVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIASIGCGVTMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374

QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 375 QAIG---GSIPAFILGSLGVIANQVQTPVEVIMGIRTSIALVPCGFMLLAFVFIWFYPLTD 431

QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221
Db 432 K--KPKSEIVVEIDNRKKVQQQLISDIT 456

RESULT 15
B64918
glucuronide permease uidB - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: B64918
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64918
A;Status: nucleic acid
A;Molecule type: DNA
A;Residues: 1-457 <BLAT>
A;Cross-references: UNIPROT:P30868; GB:AE000257; GB:U00096; NID:g1787898; PIDN:AAC74688.
A;Experimental source: strain K-12, substrain MG1655
```

```
C;Genetics:
A;Gene: uidB; gusB
C;Superfamily: melibiose carrier protein
C;Keywords: carrier protein; transmembrane protein
F;37-53/Domain: transmembrane #status predicted <TM01>
F;81-97/Domain: transmembrane #status predicted <TM02>
F;152-168/Domain: transmembrane #status predicted <TM03>
F;184-200/Domain: transmembrane #status predicted <TM04>
F;231-247/Domain: transmembrane #status predicted <TM05>
F;263-279/Domain: transmembrane #status predicted <TM06>
F;310-326/Domain: transmembrane #status predicted <TM07>
F;408-424/Domain: transmembrane #status predicted <TM08>
```

Query Match 6.8%; Score 81.5; DB 2; Length 457;

Best Local Similarity 25.2%; Pred. No. 12;

Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

```
QY 95 WDLRSILGAVVFSRVTVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIASIGCGVTMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374

QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 375 QAIG---GSIPAFILGSLGVIANQVQTPVEVIMGIRTSIALVPCGFMLLAFVFIWFYPLTD 431

QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221
Db 432 K--KPKSEIVVEIDNRKKVQQQLISDIT 456
```

Search completed: May 27, 2005, 08:52:56

Job time : 42 secs

This Page Blank (uspto)

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 08:34:52 ; Search time 176 Seconds
(without alignments)
666.285 Million cell updates/sec

Title: US-09-942-052A-728
Perfect score: 1198
Sequence: 1 MAQPLRHRSRCATPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03: *
1: uniprot_prot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1198	100.0	229	1 OIP5 HUMAN	O43482 homo sapien
2	122	10.2	218	2 Q9CXR6	Q9cxr6 mus musculus
3	115.5	9.6	204	2 Q9C2J6	Q9c2j6 mus musculus
4	113	9.4	233	1 CU45 HUMAN	Q9ny99 homo sapien
5	110.5	9.2	232	2 Q68UT5	Q68ut5 pan troglod
6	105	8.8	155	1 YCZC SCHPO	Q9p802 schizosacch
7	92	7.7	361	2 Q8S6L1	Q8s6l1 oryza sativ
8	92	7.7	695	2 Q9CH87	Q9ch87 lactococcus
9	91	7.6	628	2 Q69KE0	Q69ke0 oryza sativ
10	90.5	7.6	597	2 Q83693	Q83693 treponema p
11	89	7.4	664	1 CHFR HUMAN	Q96ep1 homo sapien
12	89	7.4	893	1 POL2 BAMMA	Q65329 barley mild
13	88.5	7.4	638	2 Q75Z26	Q75z26 ciona intes
14	88	7.3	250	1 KKA7_CAMJE	Fl4508 campylobact
15	88	7.3	493	2 Q68VH2	Q68vh2 barley mild
16	87.5	7.3	634	2 Q6CX82	Q6cx82 kluyveromyc
17	87	7.3	664	1 CHFR MOUSE	Q81013 mus musculus
18	86.5	7.2	669	2 Q97FQ5	Q97fq5 clostridium
19	86	7.2	145	2 Q839T4	Q839t4 enterococcu
20	86	7.2	331	2 Q6GXE5	Q6gxe5 cochllobolu
21	86	7.2	893	2 Q68VH4	Q68vh4 barley mild
22	85	7.1	337	1 SYW1 STRAW	Q82hu1 streptomyce
23	85	7.1	671	2 Q815K5	Q815k5 plasmodium
24	85	7.1	1018	2 Q6FKN6	Q6fkn6 candida gla
25	85	7.1	2194	2 GLSN MEDSA	Q03460 medicago sa
26	85	7.1	2194	2 Q40360	Q40360 medicago sa
27	84.5	7.1	361	2 Q8YV34	Q8yv34 anabaena sp
28	84	7.0	247	2 Q64WS0	Q64ws0 bacteroides
29	84	7.0	496	2 Q96SL8	Q96sl8 homo sapien
30	84	7.0	553	2 Q6CAQ5	Q6cag5 yarrowia li
31	84	7.0	894	2 Q68VH5	Q68vh5 barley mild

32	84	7.0	894	2	Q68VH6	Q68vh6 barley mild
33	84	7.0	1129	2	Q7UKD4	Q7ukd4 rhodospirell
34	84	7.0	1947	2	Q84SE8	Q84se8 oryza sativ
35	83.5	7.0	474	2	Q6A5C8	Q6a5c8 propionibac
36	83.5	7.0	800	2	Q7Y5U5	Q7y5u5 bacteriophac
37	83	6.9	475	2	Q7XQK6	Q7xqk6 oryza sativ
38	83	6.9	579	2	Q8BEN6	Q8ben6 callitrichi
39	83	6.9	663	2	Q97J88	Q97j88 clostridium
40	82.5	6.9	143	2	Q27769	Q27769 methanobact
41	82.5	6.9	423	2	Q87A11	Q87a11 xylella fas
42	82.5	6.9	429	2	Q757S2	Q757s2 ashbya goss
43	82.5	6.9	491	2	Q6T6X5	Q6t6x5 corydoras s
44	82.5	6.9	493	2	Q23213	Q23213 arabidopsis
45	82.5	6.9	640	2	O56294	O56294 human herpe

ALIGNMENTS

RESULT 1
OIP5_HUMAN STANDARD; PRT; 229 AA.
AC O43482; Q96BX7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Opa-interacting protein 5.
GN Name=OIP5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98125741; PubMed=9466265;
RA Williams J.M., Chen G.-C., Zhu L., Rest R.F.;
RT "Using the yeast two-hybrid system to identify human epithelial cell
proteins that bind gonococcal Opa proteins: intracellular gonococci
bind pyruvate kinase via their Opa proteins and require host pyruvate
for growth."
RL Mol. Microbiol. 27:171-186(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalish D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBUNIT: Binds outer membrane protein Opa from Neisseria
gonorrhoeae.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF025441; AAC39561.1; ALT_INIT.

DR EMBL; BC015050; AAH15050.1; -.

DR H-InvDB; HIX0012150; -.

DR MIM; 606020; -.

DR GO; GO:0005515; F:protein binding; TAS.

DR GO; GO:0007154; P:cell communication; NAS.

SQ SEQUENCE 229 AA; 24691 MW; 0EBD4006193A3106 CRC64;

Query Match 100.0%; Score 1198; DB 1; Length 229;

Best Local Similarity 100.0%; Pred. No. 1.5e-103;

Matches. 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPQLRHRSRCATPRGDFCGGTERAIDQASFTTSMEDVTQVKGSSPLGAGLCAEP 60

DB 1 MAAPQLRHRSRCATPRGDFCGGTERAIDQASFTTSMEDVTQVKGSSPLGAGLCAEP 60

QY 61 AAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPFL 120

DB 61 AAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPFL 120

QY 121 VGIEGLKSTYNLLFCGSGGIPVGFHLYSTHAALRGHFCLSDDKXWVCYLLTKAIV 180

DB 121 VGIEGLKSTYNLLFCGSGGIPVGFHLYSTHAALRGHFCLSDDKXWVCYLLTKAIV 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

DB 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 2

Q9CXK6 PRELIMINARY; PRT; 218 AA.

AC Q9CXK6;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched

DE library, clone:3110025H23 product:hypothetical protein, full insert

DE sequence.

GN Name=2610039C10Rik;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA The FANTOM Consortium;

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (KISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,

RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,

RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,

RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,

RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,

RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,

RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,

RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,

RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,

RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AK014084; BAB39147.1; -.

DR MGD; MGI:1913828; 2610039C10Rik.

KW Hypothetical protein.

SQ SEQUENCE 218 AA; 24447 MW; 55390B5BFDD8C2 CRC64;

Query Match 10.2%; Score 122; DB 2; Length 218;

Best Local Similarity 23.3%; Pred. No. 0.0046;

Matches 48; Conservative 30; Mismatches 102; Indels 26; Gaps 6;

QY 29 DOASFTTSMEDVTQVKGSSPLGAGLCAEPAGLPQLPSWLQPERCAVFOCAQCHAVLA 88

DB 15 DSSRYLRQLQWAN--MSSADALGLEKEPERKAAAEHP-----LVFLCARCRRPLG 64

QY 89 DSVHLAWDLR-SLGAVVFSRVTVNNVLEAPFLVGIEGLKSTYNLLFCGSGGIPVGFH 147

DB 65 DS--LTVVASQBDTNCILLRSVSCNVSDKBPKLKCRDEDCILEALYCTGSLSLGYV 122

QY 148 LYSTHAALAAALRGHFCLSDDKXWVCYLL-----KTKAIVNASEMDIQNVPLSEKIAE 198

DB 193 VRCPTPNLDYKRDLCFLSVETVETSLGSSBKQIVSKDKELFNLE----SRVEIKSIQ 178

QY 199 LKEKIVLTHNRLKSLMKILSEVTPDQ 224

DB 179 MEEVLTALQKRLREVEKSLSLAQPCQ 204

RESULT 3

Q9CZJ6 PRELIMINARY; PRT; 204 AA.

AC Q9CZJ6;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Mus musculus 11 days embryo whole body cDNA, RIKEN full-length

DE enriched library, clone:2700078124 product:hypothetical protein, full

DE insert sequence (2610039C10Rik protein).

GN Name=2610039C10Rik;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria, Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P. Hayaahizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayaahizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiseunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakauchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayaahizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayaahizaki Y.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Eye;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Heilef F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stratton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshikiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Eye;
RC Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK012533; BAB28302.1; -;
DR EMBL; BC079900; AAH79900.1; -;
DR MGD; MGI:1913828; 2610039C10Rik.
KW Hypothetical protein.
SQ SEQUENCE 204 AA; 22949 MW; BBD667250DF9B0F CRC64;
Query Match 9.6%; Score 115.5; DB 2; Length 204;
Best Local Similarity 25.6%; Pred. No. 0.017;
Matches 50; Conservative 29; Mismatches 97; Indels 19; Gaps 7;
QY 29 DOASFTTSMEDWTQVVKVSGPLGAGAEPAAGPQLPSWLOPERCAVQCACQACHVLA 88
Db 15 DSRYLRLQKXAN--MSSADALGLEKEPEREKAANAENP-----LVFLCARCRPLG 64
QY 89 DSVHLAWDLNR-SLGAVFVSRTNNVLEAPFLVGLGSLKSTYNNLLFCGSGIPGVGFH 147
Db 65 DS--LTWVASQEDTNCILRSVSCNVSDKEPKSKRDEDCILREALYCTGSLGYV 122
QY 148 LVSTHAALALRGHFCLSDDKMKVCYLL--KTAIVNASEMDIONVPLSEKIAELKEKIVL 205
Db 123 YRCTPKNLYDKDLFCLSVEAVESYTLGSEKQIV--SEDKELFNL---ESRVEIEKSIKQ 178
QY 206 TNRLKSLMKILSEV 220
Db 179 MEEVLTALQKLEEV 193
RESULT 4
CU45_HUMAN STANDARD; PRT; 233 AA.
AC Q9NTE9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Putative protein C21orf45.
GN Name=C21orf45;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20337674; PubMed=10773462; DOI=10.1016/S0378-1119(00)00089-5;
RA Slavov D., Hattori M., Sakaki Y., Rosenthal A., Shimizu N.,
RA Minoshima S., Kudo H., Yaspo M.-L., Ramser J., Reinhardt R.,
RA Reimer C., Clancy K., Rynditch A., Gardiner K.;
RT "Criteria for gene identification and features of genome organization:
RT analysis of 6.5 Mb of DNA sequence from human chromosome 21.";
RL Gene 247:215-232(2000).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan M., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Groomwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinaki M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- SIMILARITY: To S.pombe C970.12.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF231921; AAP72945.1; -;
 DR EMBL; BC042917; AAH42917.1; -;
 DR Genew; HGNC:1286; C21orf45.
 KW Hypothetical protein.
 SQ SEQUENCE 233 AA; 25863 MW; A07522806C4B6221 CRC64;

Query Match 9.4%; Score 113; DB 1; Length 233;
 Best Local Similarity 27.1%; Pred. No. 0.035;
 Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;
 QY 31 ASFTTSMEDTQVVGSSPLGAGAEPPAGPQPSLWQPERCAVFOCAOCHAVLADS 90
 DB 49 ASWSSMSDASV-----ADMERARL--EEAANA-----ERPLVFLCSCGRPLGDS 95
 QY 91 VHLAWDLR-SLGAVVFSRVTVNNVLEAPFLVGIEGSLKSTYNNLLFCGSCGIPVGFHLY 149
 DB 96 --LSWVASQEDTNCILLRCVNCVSDKEQKSKREKENGCVLETLCCAGCSLNLGYVVR 153
 QY 150 STHAALAARGHFLCSSLDRKVCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
 DB 154 CTPKNLDYKRDLCFSLVRAIESYVLGSSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209
 QY 208 NRLKSLMKILSE 219
 DB 210 DVLKALOMKLWE 221

RESULT 5
 Q68UT5 PRELIMINARY; PRT; 232 AA.
 AC Q68UT5;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Hypothetical protein C21orf45.
 GN Name=C21orf45;
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RG The Chimpanzee Chromosome 22 Sequencing Consortium;
 RA Watanabe H., Fujiyama A., Hattori M., Taylor T.D., Toyoda A.,
 RA Kuroki Y., Noguchi H., Benkhal A., Lehrach H., Lehrach R., Kube M.,

RA Taenzer S., Galgoczy P., Platzer M., Scharfe M., Nordsiek G.,
 RA Blocker H., Hellmann I., Khaitovich P., Paabo S., Reinhardt R.,
 RA Zheng H.J., Zhang X.L., Zhu G.F., Wang B.P., Fu G., Ren S.X.,
 RA Zhao G.P., Chen Z., Lee Y.S., Cheong J.E., Choi S.H., Wu K.M.,
 RA Liu T.T., Hsiao K.J., Tsai S.F., Kim C.G., Oota S., Kitano T.,
 RA Kohara Y., Saitou N., Park H.S., Wang S.Y., Yaspo M.L., Sakaki Y.,
 RT "DNA sequence and comparative analysis of chimpanzee chromosome 22.";
 RL Nature 429:382-388 (2004).
 DR EMBL; AL954205; CAH18576.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 232 AA; 25831 MW; 6549BBBE369245E7 CRC64;
 Query Match 9.2%; Score 110.5; DB 2; Length 232;
 Best Local Similarity 27.1%; Pred. No. 0.059;
 Matches 52; Conservative 29; Mismatches 88; Indels 23; Gaps 8;
 QY 31 ASFTTSMEDTQVVGSSPLGAGAEPPAGPQPSLWQPERCAVFOCAOCHAVLADS 90
 DB 49 ASWSSMSDASV-----ADMERARL--EEAANA-----ERPLVFLCSCGRPLGDS 94
 QY 91 VHLAWDLR-SLGAVVFSRVTVNNVLEAPFLVGIEGSLKSTYNNLLFCGSCGIPVGFHLY 149
 DB 95 --LSWVASQEDTNCILLRCVNCVSDKEQKSKREKENGCVLETLCCAGCSLNLGYVVR 152
 QY 150 STHAALAARGHFLCSSLDRKVCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
 DB 153 CTPKNLDYKRDLCFSLVRAIESYVLGSSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 208
 QY 208 NRLKSLMKILSE 219
 DB 209 DVLKALOMKLWE 220

RESULT 6
 YCZC SCHPO STANDARD; PRT; 155 AA.
 ID YCZC SCHPO
 AC Q9P802;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypothetical protein C970.12 in chromosome III.
 GN ORFNames=SPCC970.12;
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Wambutt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

Query Match 7.7%; Score 92; DB 2; Length 695;
Best Local Similarity 25.3%; Pred. No. 13;
Matches 38; Conservative 28; Mismatches 50; Indels

```
QY 101 LGAVFSRVNNVLEAPFLVGLGSLKSTYNLLFCGSGGIPVGFHLYSTHAAALALRG 160
Db 101 LGTIIF-----FYSCTPFSGAKGLSKRKPKAMMLITWGITVAY-AYSVYATIMSLNG 153
QY 161 HF-----CLSSDKWCYLLTKTKAIVNASE--NDI-----QNVPLSE-K 195
Db 154 HMGWTFWFLATLIVIMLGLHLETKAMKAGDALKDLASLPVKKAHLSKGKDVSELK 213
QY 196 IABLKEKIVLTHNRKLSMKILSEVTPDOS 225
Db 214 VGDL--LLVKNKIKIPADGLILSEALVDES 241
RESULT 9
Q69KE0 PRELIMINARY; PRT; 628 AA.
AC Q69KE0;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE Myosin heavy chain-like.
GN Names=P0023E10.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;
RT "Oryza sativa nipponbare (GA) genomic DNA, chromosome 9, PAC
clone:P0023E10.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005934; BAD36600.1; -.
DR InterPro; IPR005829; Sug transporter.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN 1.
SQ SEQUENCE 628 AA; 68281 MW; 3230B108C6061027 CRC64;
Query Match 7.6%; Score 91; DB 2; Length 628;
Best Local Similarity 23.7%; Pred. No. 14;
Matches 66; Conservative 34; Mismatches 82; Indels 96; Gaps 17;
QY 8 HRSRCATPRG-----DFCGTERAIOAQFTTSMWDTQVVKSSPLGALGA 57
Db 141 YRKAPTSPSRVARAKAQAQSGGTSSAPAVAST-----DVVVWPGSREAPSGL-A 194
QY 58 BPAAAGPQLP---SW--LQPERCAVFC-----AQCHAVLADSVH-----LAWDL 97
Db 195 SOLVAGPGSPAALVTWEELQVEMGRLLLEAGARGIGREITAEARAAASSANERADRLAHD 254
QY 98 S-----RSLGAV-----FSRVNNVLEAPFLVGLGSLKSTYNLL--F 136
Db 255 AEAREDLKQRELVAQNRQWGLEHRMSLENN-----LSEIRDSLR-VYITGLHQL 306
QY 137 CQSGI-----PVGFHLYSTHAAAL-----RGHFLCLSSDKM-----VCYLLTKTK 177
Db 307 AGECKGKSTIPANPDFFSLTSLAELAAAMEIPSKHAARTRETSNGIYTGACHVLACL 366
QY 178 AIVNASEMDIONV-----PLSEKIAELKEKIV 204
Db 367 RLVH-PELDRLILDOGAASDARKQVMEVGLGSEVL 403
RESULT 10
O83693 PRELIMINARY; PRT; 597 AA.
ID O83693;
AC O83693;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE Phosphoribosylglycinamide formyltransferase, putative.
GN OrderedLocusNames=TP0695;
OS Treponema pallidum.
```

```
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artiach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete.";
RL Science 281:375-388(1998).
DR EMBL; AE001243; AAC65662.1; -.
DR PIR; D71293; D71293.
DR TIGR; TP0695; -.
DR InterPro; IPR003135; ATP-grasp.
DR Pfam; PF02222; ATP-grasp; 1.
KW Complete proteome.
SQ SEQUENCE 597 AA; 63325 MW; 7B775117C9BFD6A6 CRC64;
Query Match 7.6%; Score 90.5; DB 2; Length 597;
Best Local Similarity 28.0%; Pred. No. 15;
Matches 51; Conservative 19; Mismatches 47; Indels 65; Gaps 12;
QY 12 CATPRGDFCGGTERAID-----QASFT-----TSMWDTQVVKSSPLGPA 53
Db 90 CALP--GHRLEATKATKTRMRACFTRARLCRPTFLEPDSFAMD-----PPGHA 140
QY 54 GLGAEPAAGPOLPSWLP--ERCAVFC--AQCHAVLADSVHLAWDLRSLSGLGAVFSRV 110
Db 141 RLCSHLHSAGLSFPLVVRPTDNMGARGCTLAQCKDTLNACAVARQFSRS----- 190
QY 111 NNVLLEAPFLVGLGSLKSTYNLLFCGSGGIPVGFHLYSTHAAALALRGHFLCLSSDKMV 170
Db 191 GRVIEE-BFIVGREPSLEG---LIFDGT-----LYVT--ALA-----DRHI 225
QY 171 CY 172
Db 226 CF 227
RESULT 11
CHFR_HUMAN STANDARD; PRT; 664 AA.
ID CHFR_HUMAN STANDARD; PRT; 664 AA.
AC Q96EP1; Q96SL3; Q9NRT4; Q9NT32; Q9NVD5;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ubiquitin ligase protein CHFR (EC 6.3.2.-) (Checkpoint with forkhead
DE and RING finger domains protein).
GN Name=CHFR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY, AND
RP VARIANT MET-580.
RX MEDLINE=20388685; PubMed=10935642; DOI=10.1038/35019108;
RA Scolnick D.M., Halazonetis T.D.;
RT "Chfr defines a mitotic stress checkpoint that delays entry into
RT metaphase.";
RL Nature 406:430-435(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Teratocarcinoma;
RX PubMed=14702039; DOI=10.1038/ngl1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
```

RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Iehii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shigatori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Iehibashi T., Yamaoka H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hottu T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki M., Watanabe T., Sugiyama A., Takenoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANT VAL-497.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gichmond J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 359-664 FROM N.A.
RC TISSUE=Testis;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RN Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION, AUTOUBIQUITINATION, AND MUTAGENESIS OF ILE-306 AND TRP-332.
RX PubMed=11807090; DOI=10.1093/jcb.200108016;
RA Kang D., Chen J., Wong J., Fang G.;
RT "The checkpoint protein Chfr is a ligase that ubiquitinates Plk1 and
RT inhibits Cdc2 at the G2 to M transition.";
RL J. Cell Biol. 156:249-259(2002).
RN [6]
RP FUNCTION, AUTOUBIQUITINATION, SUBCELLULAR LOCATION, AND DEVELOPMENTAL
RP STAGE.
RX PubMed=11912157;
RA Chaturvedi P., Sudakin V., Bobiak M.L., Fisher P.W., Mattern M.R.,
RA Jablonski S.A., Hurle M.R., Zhu Y., Yen T.J., Zhou B.-B.;
RT "Chfr regulates a mitotic stress pathway through its RING-finger
CC

RT domain with ubiquitin ligase activity.";
RL Cancer Res. 62:1797-1801(2002).
RN [7]
RP DISEASE, AND VARIANTS ARG-270; VAL-497 AND MET-580.
RX PubMed=11948416; DOI=10.1038/sj.onc.1205402;
RA Mizuno K., Osada H., Koiishi H., Tatematsu Y., Yatabe Y.,
RA Mitudomi T., Fujii Y., Takahashi T.;
RT "Aberrant hypermethylation of the Chfr prophase checkpoint gene in
RL human lung cancers.";
RL Oncogene 21:2328-2333(2002).
RN [8]
RP DISEASE.
RX PubMed=12538348; DOI=10.1093/carcin/24.1.47;
RA Corn P.G., Summers M.K., Fogt F., Virmani A.K., Gazdar A.F.,
RA Halazonetis T.D., El-Deiry W.S.;
RT "Frequent hypermethylation of the 5' CpG island of the mitotic stress
RT checkpoint gene Chfr in colorectal and non-small cell lung cancer.";
RL Carcinogenesis 24:47-51(2003).
RN [9]
RP DISEASE.
RX PubMed=12810945; DOI=10.1073/pnas.1337066100;
RA Toyota M., Sasaki Y., Satoh A., Ogi K., Kikuchi T., Suzuki H.,
RA Mita H., Tanaka N., Itoh F., Issa J.-P.J., Jair K.-W., Schuebel K.E.,
RA Imai K., Tokino T.;
RT "Epigenetic inactivation of Chfr in human tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7818-7823(2003).
RN [10]
RP DISEASE.
RX PubMed=14695171;
RA Satoh A., Toyota M., Itoh F., Sasaki Y., Suzuki H., Ogi K.,
RA Kikuchi T., Mita H., Yamashita T., Kojima T., Kusano M., Fujita M.,
RA Hosokawa M., Endo T., Tokino T., Imai K.;
RT "Epigenetic inactivation of Chfr and sensitivity to microtubule
RT inhibitors in gastric cancer.";
RL Cancer Res. 63:8606-8613(2003).
RN [11]
RP DISEASE, AND VARIANTS NSCLC LEU-166; PRO-202 AND SER-536.
RX PubMed=14612512;
RA Mariatos G., Bothos J., Zacharatos P., Summers M.K., Scolnick D.M.,
RA Kittas C., Halazonetis T.D., Gorgoulis V.G.;
RT "Inactivating mutations targeting the Chfr mitotic checkpoint gene in
RT human lung cancer.";
RL Cancer Res. 63:7185-7189(2003).
RN [12]
RP PHOSPHORYLATION, AND MUTAGENESIS OF THR-39 AND SER-205.
RX PubMed=14638868;
RA Shivelman E.;
RT "Promotion of mitosis by activated protein kinase B after DNA damage
RT involves polo-like kinase 1 and checkpoint protein Chfr.";
RL Mol. Cancer Res. 1:959-969(2003).
RN [13]
RP FUNCTION, INTERACTION WITH UBE2V2, AND PHOSPHORYLATION.
RX PubMed=14562038; DOI=10.1038/sj.onc.1206831;
RA Bothos J., Summers M.K., Veneris M., Scolnick D.M., Halazonetis T.D.;
RT "The Chfr mitotic checkpoint protein functions with Ubc13-Mms2 to form
RT Lys63-linked polyubiquitin chains.";
RL Oncogene 22:7101-7107(2003).
RN [14]
RP FUNCTION.
RX PubMed=14694445; DOI=10.1002/mc.10161;
RA Erson A.E., Petty E.M.;
RT "Chfr-associated early G2/M checkpoint defects in breast cancer
RT cells.";
RL Mol. Carcinog. 39:26-33(2004).
RN [15]
RP X-RAY CRYSTALLOGRAPHY (2.0 ÅNGSTRÖMS) OF 14-128.
RX PubMed=12121644; DOI=10.1016/S0969-2126(02)00776-1;
RA Stavridi E.S., Huyen Y., Loreto I.R., Scolnick D.M., Halazonetis T.D.,
RA Pavlath N.P., Jeffrey P.D.;
RT "Crystal structure of the FHA domain of the Chfr mitotic checkpoint
RT protein and its complex with tungstate.";
RL Structure 10:891-899(2002).
CC -!- FUNCTION: E3 ubiquitin-protein ligase required to transiently

CC arrest cells in early prophase when they are exposed to
 CC microtubule poisons. Acts in early prophase before chromosome
 CC condense, when the centrosome move apart from each other along the
 CC periphery of the nucleus. Probably promotes the formation of Lys-
 CC 63-linked polyubiquitin chains and functions with the specific
 CC ubiquitin-conjugating UBC13-WMS2 (UBE2N-UBE2V2) heterodimer.
 CC Substrates that are polyubiquitinated at Lys-63 are usually not
 CC targeted for degradation, but are rather involved in signaling
 CC cellular stress. This suggests that it may be involved in
 CC signaling the presence of mitotic stress caused by microtubule
 CC poisons.
 CC -|- PATHWAY: Ubiquitin conjugation; third step.
 CC -|- SUBCELLULAR LOCATION: Nuclear.
 CC -|- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q96EP1-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q96EP1-2; Sequence=VSP_009349;
 CC Name=3;
 CC IsoId=Q96EP1-3; Sequence=VSP_009350;
 CC Note=No experimental confirmation available;
 CC -|- TISSUE SPECIFICITY: Ubiquitous.
 CC -|- DEVELOPMENTAL STAGE: Weakly expressed in G1 phase, and highly
 CC expressed during S phase.
 CC -|- DOMAIN: The FHA domain may be required to interact with
 CC phosphorylated proteins.
 CC -|- PTM: Autoubiquitinated in vitro.

Query Match 7.4%; Score 89; DB 1; Length 664;
 Best Local Similarity 23.0%; Pred. No. 23;
 Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
 QY 8 HRSRCATPRGDFC-----GGTERAIDQA-----SFTTSMEDTQVVKSSPL-----50
 Db 434 YRQAALQPH-----CPAPEGEPQALGDAPSTSVSLTTAVQDYVCPQGSALCTCCFQ 490
 QY 51 ---GPAGLGAEEPAAGPQLPQLQPERCAVFOCAOCHAVLADSVHLAWDLR-----99
 Db 491 PMPDRAEREQDPRAVQ-----QCAVC-----LQPFCHLYWCCTGCGYGLA 534
 QY 100 -----SLGAVFSRVNNVNVLEAPFLVIGIEGSLKSTY-NLLFCGSCGIPVGFHLYSTHA 153
 Db 535 PFCELNGLDKCLDGLVNNNSYSDILKNYLAT-RGLTWKNML-----TES 578
 QY 154 ALAALRGHFCIS-----SDKWVCLLTKTAIVNASEMDIQNVPLSE 194
 Db 579 LVALQRGVLLSDYRVGTGTVLCYCGLRSFRLTYQYFQNPASE 624

RESULT 12
 POL2 BAMWA
 ID POL2 BAMWA STANDARD; PRT; 893 AA.
 AC Q65329;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Genome polyprotein 2 [Contains: Helper component proteinase
 DE (EC 3.4.22.45) (HC-pro); 70 kDa protein].
 OS Barley mild mosaic virus (strain ASL) (BAMMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Bymovirus.
 OC NCBI_TaxId=103899;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Timpe U., Kuehne T.;
 RT "The complete nucleotide sequence of RNA2 of barley mild mosaic virus
 RT (BAMMV).";
 RL Eur. J. Plant Pathol. 100:233-241(1994).
 CC -|- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|-Gly bond at its own C-
 CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-|-Gly, in the
 CC processing of the potyviral polyprotein.
 CC -|- SIMILARITY: Contains 1 peptidase C6 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X75933; CAA53537.1; --
 CC MEROPS: C06.001; --
 CC InterPro: IPR001337; TMV coat.
 CC Pfam: PF0721; TMV coat; 1.
 CC Hydrolase; Polyprotein; Thiol protease.
 FT CHAIN 1 229 Helper component proteinase (Potential).
 FT CHAIN 230 893 70 kDa protein.
 FT ACT_SITE 117 117 Potential.
 FT ACT_SITE 189 189 Potential.
 FT SITE 229 230 Cleavage (by HC-PRO protease)
 FT SITE (Potential).
 SQ SEQUENCE 893 AA; 98328 MW; D6EB74D9D6CF6407 CRC64;
 Query Match 7.4%; Score 89; DB 1; Length 893;
 Best Local Similarity 22.3%; Pred. No. 33;
 Matches 79; Conservative 36; Mismatches 105; Indels 134; Gaps 18;
 QY 4 QPLRHRSRCATPRGD---FCGTERAIDQ-ASFTTSMEW-----DTQVVKSSPLGPA 53
 Db 13 QVLR---RFSIPTSGDRLLVNSSTDQPIGLFGAFDTSLQTLQVNDPEVLKQKSI-PT 68
 QY 54 GLG-AEPEAAGPOLPSW-----LQPERCAVFOCAOCHAVL-----87
 Db 69 HLDVASVLEASPRSPFWFLTNSFCTFGSIHAQNLQAFATAEFKSGFCYMNLLIPLSPD 128
 QY 88 -----ADSVH-LAWDLRSLGAVFSRVNNVLEA-----PFLV-----121
 Db 129 ITDAHADSFRGFVEQLPDTLGLAYPSLSMVLNMLHAATRFPEIVASPIPTIAFDESLOF 188
 QY 122 -----GTEGS---LKG-STYNLLFCGSCGIPVGFHLYSTHA-----LAALRGHFCIS 166
 Db 189 HVTDRGVPGMNNILKACRVYELLSLAADGICFCYVLPVGAAPQYFVKKSMDF--TS 246
 QY 167 DMVCCYLLTKTAIVNASEMD-----IQN-----189
 Db 247 DRFVEFLAQDLLASALEQDVATHDARDALLSALQNAVYNNVVRERRFPNGHDPSIVWL 306
 QY 190 -----VPLSEKTAELKEKIV-----LTHNRLKSLMKILSEVTPOOSKPE 229
 Db 307 NLSEAPISEKLTELKRYLLVGHRSDDTDITHNHGHVFEVLKTMVSVQFSKTTN 360

RESULT 13
 Q75R26
 ID Q75R26 PRELIMINARY; PRT; 638 AA.
 AC Q75R26;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Vesicular GABA transporter.
 GN Name=Ci-vGAT;
 OS Ciona intestinalis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Clonidae; Ciona.
 OX NCBI_TaxId=7719;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshida R., Sakurai S., Horie T., Kawakami I., Tsuda M., Kusakabe T.;
 RT "Identification of neuron-specific promoters in Ciona intestinalis.";
 RL Genesis 39:130-140(2004).
 DR EMBL: AB158401; BAD06308.1; --
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005279; F:amino acid-polyamine transporter activity; IEA.

```

Best Local Similarity 28.7%; Pred. No. 8.2;
Matches 33; Conservative 18; Mismatches 46; Indels 18; Gaps 5;

QY 105 VFGRVTNNVLEAPFLVGLGIEGSLKSTYNLLFCGSCGIPGVFHLXSTHAALALRG-HFC 163
Db 40 IFSKTTYSVKREAEEMMWLSDKLKVPDV-----IEYGVREHSEYLMSELRGHID 90

QY 164 LSSDKMVCYLLTKTKAIVNA-----SEMDIQNVLPSEKI-AELKEKIVLTHNRLKSL 213
Db 91 CFIDHPIKVI--ECLVNALHQLQADIRNCPFSKIDVRLKELKYLDDNRIADI 142

RESULT 15
Q68VH2 PRELIMINARY; PRT; 493 AA.
ID Q68VH2;
AC Q68VH2;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Polyprotein.
OS Barley mild mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Bymovirus.
OX NCBI_TaxID=12466;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15290374;
RA Kanyuka K., McGrann G., Alhudaib K., Hariti D., Adams M.J.;
RT "Biological and sequence analysis of a novel European isolate of
RT Barley mild mosaic virus that overcomes the barley rym5 resistance
RT gene.";
RL Arch. Virol. 149:1469-1480(2004).
DR EMBL; AJ544273; CAD66665.1; -.
DR InterPro; IPR001337; TMV_coat.
DR Pfam; PF00721; TMV_coat; 1.
KW Polyprotein.
FT CHAIN 1 263 P1 protein.
FT CHAIN 263 493 P2 protein.
SQ SEQUENCE 493 AA; 54396 MW; BDD2IEDA85A9E22 CRC64;

Query Match 7.3%; Score 88; DB 2; Length 493;
Best Local Similarity 22.3%; Pred. No. 19;
Matches 79; Conservative 36; Mismatches 105; Indels 134; Gaps 18;

QY 4 QPLHRSRCATPRGD---FCGGERAIDQ-ASFTTSMEW-----DTQVVGKSSPLGPA 53
Db 13 QVLR---RPSIPTSGRLIVNSDTPIGLFGADTSLQTLSTQVNDPEVLKQKSN-PT 68

QY 54 GLG-AEPPAAGPOLPSW-----LQPERCAVQCAQCHAVL----- 87
Db 69 HLDVASVLEASPRSPFWVFLTNSFCTFGGSIHAQNLQAFATSEKSGFCYMNLLPLSFD 128

QY 88 -----ADSVH-LAWDLRSIGAVFVRVTNNVLEA---PFLV----- 121
Db 129 IIDAHAADSPRGFVEQPDLTGLGAVPSVLMVNLVNLHAATRPFEIVASPIPTAFDAESLQF 188

QY 122 -----GIEGS---LKG-STYNLLFCGSCGIPGVFHLXSTHAA-----LAALRGHFLCSS 166
Db 189 HVTDKGVPCWNNLKACRYVELLSLAADIGCEYMLYPVCAAPQYFWKKSMDHF--TS 246

QY 167 DRMVCYLLTKTKAIVNAEMD-----IQN----- 189
Db 247 DRSVEFLAQDILLASALEQDYATHDARDALLSALQAGYTNVVARRRFPNGHDPISVWL 306

QY 190 -----VPLSEKIAELKEKIV-----LTHNRLKSLMKILSEVTPDQSKPEN 229
Db 307 NLSEAPISEKLTRELKRYLLVGHRSDDTDITNHHQHFVFLKTMVSQVFSKTTN 360

Search completed: May 27, 2005, 08:52:10
Job time : 179 secs

```

BEST AVAILABLE COPY

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 08:34:07 ; Search time 161 Seconds
(without alignments)
550.113 Million cell updates/sec

Title: US-09-942-052a-728
Perfect score: 1198
Sequence: 1 MAAQPLRHRSRCATPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	4	AAM40223 Human pol
2	1198	100.0	229	5	ABG34856 Human can
3	1198	100.0	231	4	AAM42009 Human pol
4	138.5	11.6	164	5	ABG34855 Human cDN
5	113	9.4	233	4	AAM93668 Human pol
6	113	9.4	233	4	AAM93290 Human pro
7	113	9.4	233	7	ADJ69163 Human hea
8	113	9.4	233	8	ADJ31517 Human pro
9	92	7.7	695	5	ABBS4167 Lactococc
10	92	7.7	695	8	ADS29356 Bacterial
11	90	7.5	269	4	AAO08972 Human pol
12	89	7.4	572	8	ABM83354 Human dia
13	89	7.4	623	4	AAB93182 Human pro
14	89	7.4	623	5	ABB97233 Novel hum
15	89	7.4	652	4	AAB93168 Human pro
16	89	7.4	664	4	AAB83843 Amino aci
17	89	7.4	664	4	AAB20219 Human Chf
18	88.5	7.4	388	8	ADQ08702 Clona int
19	87.5	7.3	1245	7	ADC32818 Human nov
20	86.5	7.2	470	8	ABM83305 Human dia
21	86.5	7.2	669	6	ABU23878 Protein e
22	86	7.2	158	7	ADH8711 Enterococ
23	84.5	7.1	485	5	ABP51336 Human MDD
24	84	7.0	496	4	AAB95788 Human pro
25	84	7.0	496	7	ADC31196 Human nov

26	83.5	7.0	479	4	AAU39678	Aau39678 Propionib
27	83.5	7.0	479	6	ABM36197	Abm36197 Propionib
28	83	6.9	598	2	AAW59461	Aaw59461 Microbial
29	83	6.9	598	2	AAW59463	Aaw59463 Microbial
30	83	6.9	598	2	AAW59457	Aaw59457 Microbial
31	83	6.9	598	2	AAW59458	Aaw59458 Microbial
32	83	6.9	598	2	AAW59459	Aaw59459 Microbial
33	83	6.9	598	2	AAW59462	Aaw59462 Microbial
34	83	6.9	598	2	AAW59460	Aaw59460 Microbial
35	83	6.9	598	2	AAW59464	Aaw59464 Microbial
36	83	6.9	599	2	AAW59496	Aaw59496 Microbial
37	83	6.9	606	1	AAW90110	Aap90110 Polypteti
38	83	6.9	607	2	AAW59456	Aaw59456 Microbial
39	83	6.9	607	2	AAW59465	Aaw59465 Microbial
40	83	6.9	663	6	ABU23749	Abu23749 Protein e
41	82.5	6.9	325	3	AAW90242	Aay90242 Human cyc
42	82.5	6.9	493	8	ADN72405	Adn72405 Thale cre
43	82.5	6.9	1043	8	ADN27140	Adn27140 Bacterial
44	82	6.8	2208	8	ADN72815	Adn72815 Thale cre
45	81.5	6.8	456	1	AAW91895	Aap1895 Protein s

ALIGNMENTS

RESULT 1
AAM40223
ID AAM40223 standard; protein; 229 AA.

XX AAM40223;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3368.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI59379.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Example 5; SEQ ID NO 3368; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotide
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX

SQ Sequence 229 AA;

Query Match 100.0%; Score 1198; DB 4; Length 229;
Best Local Similarity 100.0%; Pred. No. 8.4e-127;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAAGLAEPP 60
DB 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAAGLAEPP 60
QY 61 AAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPFL 120
DB 61 AAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPFL 120
QY 121 VGIEGLKSGSTYNLLFCGSCGIPVGFHLYSTHAALAAALRGHFCCLSSDKMVCYLLKTKAIV 180
DB 121 VGIEGLKSGSTYNLLFCGSCGIPVGFHLYSTHAALAAALRGHFCCLSSDKMVCYLLKTKAIV 180
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

-RESULT 2

ABG34856
ID ABG34856 standard; protein; 229 AA.

XX ABG34856;

AC
XX
DT 15-JUL-2002 (first entry)

XX Human cancer related protein encoded by cDNA 85P1B3.

XX Human; cytostatic; 85P1B3; cancer; immunogen; chromosome 15q14.

XX Homo sapiens.

XX WO200218578-A2.

XX 07-MAR-2002.

XX 28-AUG-2001; 2001WO-US026838.

XX 28-AUG-2000; 2000US-0228432P.

XX (AGEN-) AGENSYS INC.

XX Raitano AB, Faris M, Hubert RS, Afar D, Ge W, Challita-Bid P;
XX Jakobovits A;

XX WPI; 2002-382963/41.
XX N-PSDB; ABK70506.

XX Composition for modulating the status of 85P1B3 protein or a molecule
PT comprising a substance e.g. antibody specific to, nucleic acid encoding,
PT or fibrozyme of 85P1B3.

XX Claim 34; Fig 2; 201pp; English.

PS The invention relates to a composition comprising a substance that
XX modulate the status of 85P1B3, where the status of a cell expresses
CC 85P1B3 gene product is modulated. Also included are a composition
CC comprising a peptide region of 5 amino acids of the 85P1B3 protein, in
CC any whole number increment up to 229 that includes an aa position
CC selected from an aa position having a value greater than 0.5 in the
CC hydrophilicity profile, an aa position having a value less than 0.5 in the
CC the hydrophobicity profile, an aa position having a value greater than
CC 0.5 in the percent accessible residue profile, an aa position having a
CC value greater than 0.5 in the average flexibility profile, or an aa
CC position having a value greater than 0.5 in the beta-turn profile; a
CC polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous
CC residues of the 85P1B3 protein; a recombinant protein comprising the
CC antigen-binding region of a monoclonal antibody; a non-human transgenic
CC animal that produces an antibody that binds to the 85P1B3 protein; a
CC monoclonal antibody (MAb) that produces antibody specific to the protein; a single chain
CC heavy and monoclonal antibodies specific to the protein; a vector
CC comprising a polynucleotide that encodes the MAb; inhibiting growth of
CC cancer cells or treating a patient who bears cancer cells that expresses
CC the protein, by administering the protein, antibody, polynucleotide
CC encoding the protein, antisense polynucleotide to the polynucleotide,
CC ribozyme that cleaves the polynucleotide and T cells that specifically
CC recognize the protein; and generating a mammalian immune response
CC directed to the protein exposing cells of the mammal's immune system to
CC an immunogenic portion of the protein or polynucleotide. The composition,
CC which comprises an antibody specific to the protein, is useful for
CC delivering a cytotoxic agent to a cell that expresses the protein by
CC providing a cytotoxic agent conjugated to antibody and exposing the cell
CC to the antibody-agent conjugate. The methods are useful for inhibiting
CC growth of cancer cells or treating a patient who bears cancer cells that
CC expresses the protein, for generating a mammalian immune response
CC directed to the protein, for detecting the presence of the protein or
CC polynucleotide in a biological sample in a patient who has or who is
CC suspected of having cancer and for monitoring 85P1B3 in a biological
CC sample from a patient who has or who is suspected of having cancer. The
CC gene for 85P1B3 is located on human chromosome 15q14. The present
CC sequence is the 85P1B3 protein
XX

SQ Sequence 229 AA;

Query Match 100.0%; Score 1198; DB 5; Length 229;
Best Local Similarity 100.0%; Pred. No. 8.4e-127;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAAGLAEPP 60
DB 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAAGLAEPP 60
QY 61 AAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPFL 120
DB 61 AAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPFL 120
QY 121 VGIEGLKSGSTYNLLFCGSCGIPVGFHLYSTHAALAAALRGHFCCLSSDKMVCYLLKTKAIV 180
DB 121 VGIEGLKSGSTYNLLFCGSCGIPVGFHLYSTHAALAAALRGHFCCLSSDKMVCYLLKTKAIV 180
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 3

AAM42009
ID AAM42009 standard; protein; 231 AA.

XX AAM42009;

XX 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 6940.
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI61165.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
PT
XX
XX Example 2; SEQ ID NO 6940; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAI38642-AAI42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 231 AA;
XX
XX Query Match 100.0%; Score 1198; DB 4; Length 231;
XX Best Local Similarity 100.0%; Pred. No. 8.5e-127;
XX Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MAAQPLHRSRCATPPRGDFCGGTERAIDQASFTTSMENDTQVKGSSPLGPAAGLGAERP 60
DB 3 MAAQPLHRSRCATPPRGDFCGGTERAIDQASFTTSMENDTQVKGSSPLGPAAGLGAERP 62
QY 61 AAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSGLGVVPSRVTVNNVLEAPPL 120
DB 63 AAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSGLGVVPSRVTVNNVLEAPPL 122
QY 121 VGEISLKGSTYNLLFCGSGIPVGFHLYSTHAALARGHFCFLSSDKMVCYLLTKTAIV 180

Db 123 VGEISLKGSTYNLLFCGSGIPVGFHLYSTHAALARGHFCFLSSDKMVCYLLTKTAIV 182
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 183 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 231
RESULT 4
ABG34855
ID ABG34855 standard; protein; 164 AA.
XX
AC ABG34855;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cDNA 85P1B3 splice variant, open reading frame #3.
XX
KW Human; cytostatic; 85P1B3; cancer; immunogen; chromosome 15q14.
XX
OS Homo sapiens.
XX
PN WO200218578-A2.
XX
PD 07-MAR-2002.
XX
PF 28-AUG-2001; 2001WO-US026838.
XX
PR 28-AUG-2000; 2000US-0228432P.
XX
XX (AGEN-) AGENSYS INC.
XX
PI Raitano AB, Faris M, Hubert RS, Afar D, Ge W, Challita-Bid P;
PI Jakobovits A;
XX
DR WPI; 2002-382963/41.
DR N-PSDB; ABK70504.
XX
XX Composition for modulating the status of 85P1B3 protein or a molecule
XX comprising a substance e.g. antibody specific to, nucleic acid encoding,
XX or ribozyme of 85P1B3.
XX
XX Example 38; Page 124; 201pp; English.
XX
XX The invention relates to a composition comprising a substance that
XX modulate the status of 85P1B3, where the status of a cell expresses
XX 85P1B3 gene product is modulated. Also included are a composition
XX comprising a peptide region of 5 amino acids of the 85P1B3 protein, in
XX any whole number increment up to 229 that includes an aa position
XX selected from an aa position having a value greater than 0.5 in the
XX hydrophilicity profile, an aa position having a value less than 0.5 in
XX the hydrophobicity profile, an aa position having a value greater than
XX 0.5 in the percent accessible residue profile, an aa position having a
XX value greater than 0.5 in the average flexibility profile, or an aa
XX position having a value greater than 0.5 in the beta-turn profile; a
XX polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous
XX residues of the 85P1B3 protein; a recombinant protein comprising the
XX antigen-binding region of a monoclonal antibody; a non-human transgenic
XX animal that produces an antibody that binds to the 85P1B3 protein; a
XX hybridoma that produces antibody specific to the protein; a single chain
XX monoclonal antibody (MAB) that comprises the variable domains of the
XX heavy and monoclonal antibodies specific to the protein; a vector
XX comprising a polynucleotide that encodes the MAB; inhibiting growth of
XX cancer cells or treating a patient who bears cancer cells that expresses
XX the protein, by administering the protein, antibody, polynucleotide
XX encoding the protein, antisense polynucleotide to the polynucleotide,
XX ribozyme that cleaves the polynucleotide and T cells that specifically
XX recognize the protein; and generating a mammalian immune response
XX directed to the protein exposing cells of the mammal's immune system to
XX an immunogenic portion of the protein or polynucleotide. The composition,
XX which comprises an antibody specific to the protein, is useful for
XX delivering a cytotoxic agent to a cell that expresses the protein by
XX providing a cytotoxic agent conjugated to antibody and exposing the cell
XX to the antibody-agent conjugate. The methods are useful for inhibiting

CC growth of cancer cells or treating a patient who bears cancer cells that
 CC expresses the protein, for generating a mammalian immune response
 CC directed to the protein, for detecting the presence of the protein or
 CC polynucleotide in a biological sample in a patient who has or who is
 CC suspected of having cancer and for monitoring 85p1b3 in a biological
 CC sample from a patient who has or who is suspected of having cancer. The
 CC gene for 85p1b3 is located on human chromosome 15q14. The present
 CC sequence is the 85p1b3 protein
 . XX
 SQ Sequence 164 AA;

Query Match 11.6%; Score 138.5; DB 5; Length 164;
 Best Local Similarity 82.1%; Pred. No. 7.5e-07;
 Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;
 QY 162 FCLSSDKWVCYLKTKAIVNASRMDIQNVPLSEKIAELK 200
 DB 2 FFLSS-----YLDKTKAIVNASRMDIQNVPLSEKIAEVK 35

RESULT 5
 AAM93668
 ID AAM93668 standard; protein; 233 AA.
 XX AC AAM93668;
 XX 06-NOV-2001 (first entry)
 DT Human polypeptide, SEQ ID NO: 3550.
 DE Human; full length cDNA; cDNA synthesis; oligo-capping.
 KW Homo sapiens.
 OS Homo sapiens.
 PN EP1130094-A2.
 XX 05-SEP-2001.
 XX 07-JUL-2000; 2000EP-00114089.
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR N-PSDB; AAK94603.
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 PS Claim 8; SEQ ID NO 3550; 1380pp + Sequence Listing; English.
 CC The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesising the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a polypeptide encoded by a full length
 CC human cDNA of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in CD-ROM
 CC format directly from EPO
 XX Sequence 233 AA;
 SQ

Query Match 9.4%; Score 113; DB 4; Length 233;

Best Local Similarity 27.1%; Pred. No. 0.001;
 Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;
 QY 31 ASFTTSMEDTGVVKGSSPLGAGLGAEPAPQPLPSWLPQPCAVFOCAQCHAVLADS 90
 DB 49 ASMWSSMSDASV----ADMERAAQL--EEENAAAE-----ERPLVFLCSCGCRRLGDS 95
 QY 91 VHLAWDLRS-SLGAVVFSRVTVNNVLEAPFLVGIEGSLKGVTVNLLFCGSCGIPVGFHLY 149
 DB 96 --LSWVASQEDTNCILLRCVSCNVSDVDEKQLSKREKNGCVLETLCAGCCSLNLYVYR 153
 QY 150 STHAALALRGHFLCSSLDRMVCYLL--KTKAIVNASRMDIQNVPLSEKIAELKELVLT 207
 DB 154 CTPKNLDYKRDLFCLSVSAIESYVLGSSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209
 QY 208 NPLKSLMKILSE 219
 DB 210 DVLKALQMKLWE 221

RESULT 6
 AAG93290
 ID AAG93290 standard; protein; 233 AA.
 XX AC AAG93290;
 XX 13-SEP-2001 (first entry)
 DT Human protein HP10650.
 DE Human; gene therapy; tumour.
 KW Homo sapiens.
 OS Homo sapiens.
 PN WO200142302-A1.
 XX 14-JUN-2001.
 XX 06-DEC-2000; 2000WO-JP008631.
 XX 06-DEC-1999; 99JP-00346863.
 PR 06-DEC-1999; 99JP-00346864.
 PR 08-FEB-2000; 2000JP-00031062.
 PR 10-FEB-2000; 2000JP-00034090.
 PR 10-FEB-2000; 2000JP-00034091.
 PR 14-FEB-2000; 2000JP-00035829.
 PR 14-FEB-2000; 2000JP-00035899.
 PR 14-MAR-2000; 2000JP-00071161.
 PR 30-MAY-2000; 2000JP-00160851.
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX Kato S, Eguchi C, Saeki M;
 PI WPI; 2001-381646/40.
 DR N-PSDB; AAK68575.
 XX Human protein originated from tumor cell line, applicable as drug,
 PT reagent for studying intracellular protein networks and protein source
 PT for drug screening, also encoded cDNA for gene diagnosis and gene
 PT therapy.
 PS Claim 1; Page 303-304; 471pp; Japanese.
 CC The present sequence is a human protein. The human protein, preferably
 CC originated from tumour cell line, is applicable as a drug, a reagent for
 CC studying intracellular protein networks and a protein source for
 CC screening proteins for binding low molecular weight drugs. The human
 CC protein coding sequence is useful for gene diagnosis and gene therapy,
 CC expression vectors and transformant cells for detection of ligands and
 CC receptors
 XX Sequence 233 AA;
 SQ

Query Match 9.4%; Score 113; DB 4; Length 233;
 Best Local Similarity 27.1%; Pred. No. 0.001;
 Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTTMEWDTQVVGSSPLGAGAEPAAGPOLPSWLPQPCAVFOCAQCHAVLADS 90
 Db 49 ASWSSMSSEDAV-----ADMERACL--EEEAARAE-----ERPLVFLSCGCRPLGDS 95

QY 91 VHAWDLRS-SLGAVVFSRVNNVLEAPFLVIGSLKSGSTYNLLFCGSCGIPVGFHLY 149
 Db 96 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKENGCVLETLCCAGCSNLGVVYR 153

QY 150 STHAALALRGHCLSSDKMVCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
 Db 154 CTPKNLDYKRDPLFCLSVAEIESTVLSGSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209

QY 208 NRLKSLMKILSE 219
 Db 210 DVLKALQMKLWE 221

RESULT 7
 ADJ69163
 ID ADJ69163 standard; protein; 233 AA.
 XX AC ADJ69163;
 XX DT 06-MAY-2004 (first entry)
 XX DE Human heat mitochondrial protein as a therapeutic target SeqID969.
 XX KW mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX OS Homo sapiens.
 XX PN WO2003087768-A2.
 XX PD 23-OCT-2003.
 XX PF 04-APR-2003; 2003WO-US010870.
 XX PR 12-APR-2002; 2002US-0372843P.
 XX PR 17-JUN-2002; 2002US-0389967P.
 XX PR 20-SEP-2002; 2002US-0412418P.
 XX PA (MITO-) MITOKOR.
 XX PA (BUCK-) BUCK INST AGE RES.
 XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 XX WPI; 2003-845369/78.
 XX PT Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX Claim 1; SEQ ID NO 969; 180pp; English.
 XX CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that

CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nootropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 XX Sequence 233 AA;

Query Match 9.4%; Score 113; DB 7; Length 233;
 Best Local Similarity 27.1%; Pred. No. 0.001;
 Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTTMEWDTQVVGSSPLGAGAEPAAGPOLPSWLPQPCAVFOCAQCHAVLADS 90
 Db 49 ASWSSMSSEDAV-----ADMERACL--EEEAARAE-----ERPLVFLSCGCRPLGDS 95

QY 91 VHAWDLRS-SLGAVVFSRVNNVLEAPFLVIGSLKSGSTYNLLFCGSCGIPVGFHLY 149
 Db 96 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKENGCVLETLCCAGCSNLGVVYR 153

QY 150 STHAALALRGHCLSSDKMVCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
 Db 154 CTPKNLDYKRDPLFCLSVAEIESTVLSGSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209

QY 208 NRLKSLMKILSE 219
 Db 210 DVLKALQMKLWE 221

RESULT 8
 ADL31517
 ID ADL31517 standard; protein; 233 AA.
 XX AC ADL31517;
 XX DT 20-MAY-2004 (first entry)
 XX DE Human protein encoded by a full length cDNA clone SeqID 3550.
 XX KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method.
 XX OS Homo sapiens.
 XX PN EP1396543-A2.
 XX PD 10-MAR-2004.
 XX PF 07-JUL-2000; 2003EP-00025638.
 XX PR 08-JUL-1999; 99JP-00194486.
 XX PR 11-JAN-2000; 2000JP-00118774.
 XX PR 02-MAY-2000; 2000JP-00183865.
 XX PR 07-JUL-2000; 2000EP-00114089.
 XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX PI Ota T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2004-204755/20.
 XX DR N-PSDB; ADL31516.
 XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX Example 1; SEQ ID NO 3550; 1340pp; English.
 XX CC This invention relates to a novel primers useful for synthesising full

CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.
XX
SQ Sequence 233 AA;

Query Match 9.4%; Score 113; DB 8; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.001;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;
QY 31 ASPTTSMEDTQVVGSSPLGAGLGERPAGPQLPSWLOPERCAVFOCAOCHAVLADS 90
DB 49 ASMWSSNSDASV-----ADMERAGL--EEFAAAE-----ERPLVFLCSGCRRLGDS 95
QY 91 VHLAWLSR-SLGAVVFSRVTNVNVLEAPFLVGIEGSLKSTYNLLFCGSCGIPVGFHLY 149
DB 96 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKENGCVLETLCAGCSLNLGVYR 153
QY 150 STHAALALRGHFLSSDKMVCVLL--KTKAIVNASEMDIQNVPLSEKTAELKEKIVLTH 207
DB 154 CTPKNLDYKRDLFCLSVAEIESTYLGSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209
QY 208 NRLKSLMKILSE 219
DB 210 DVLKALQMKLWE 221

RESULT 9
ID ABB54167 standard; protein; 695 AA.
XX ABB54167;
XX
XX 29-AUG-2003 (revised)
DT 16-MAY-2002 (first entry)
DE Lactococcus lactis protein yihC.
XX
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
XX Lactococcus lactis; IL1403.
XX
XX FR2807446-A1.
XX
XX 12-OCT-2001.
XX
XX 11-APR-2000; 2000FR-00004630.
XX
XX 11-APR-2000; 2000FR-00004630.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
XX WPI; 2002-043418/06.
XX
XX New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species.
XX
XX Claim 6; SEQ ID NO 869; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
CC acid sequence is useful in the detection and/or amplification of nucleic
CC acid sequence, particularly to identify Lactococcus lactis or related
CC species. The proteins of the invention are useful for the biosynthesis or
CC biodegradation of a composition of interest. The invention helps research
CC in lactic bacteria, particularly useful in the production of yogurt and

CC cheese. Note: The sequence data for this patent is based on equivalent
CC patent WO200177334 (published 18-OCT-2001) which is available in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 695 AA;

Query Match 7.7%; Score 92; DB 5; Length 695;
Best Local Similarity 25.3%; Pred. No. 1.3;
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
QY 101 LGAVVFSRVTNVNVLEAPFLVGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRG 160
DB 101 LGTIIF-----FYSGTFFSGAKGELKSRKPMWMLITWGITVAY-AYSVYATINSLNG 153
QY 161 HF-----CLSSDKMVCVLLKTKAIVNASE--MDI-----QNVPLSE-K 195
DB 154 HGMGNFWFELATLIVIMLIGHLIENKAIMGADKDLASLVPKKAHLKSGKDVSELK 213
QY 196 IAELEKEKIVLTHNRLKSLMKILSEVTPDOS 225
DB 214 VGDL--LLVKEKEKIPADGLIILSEALVDES 241

RESULT 10
ADS29356
ID ADS29356 standard; protein; 695 AA.
XX ADS29356;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #18389.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 18389; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant

PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42006.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 572 AA;

Query Match 7.4%; Score 89; DB 8; Length 572;
Best Local Similarity 23.0%; Pred. No. 2.1;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPPRGDFC-----GTERAIDQA-----SFTTSMEDTQVVGSSPL----- 50
Db 342 YRRQAQPPH---CPAPEGEPGAPQALGDAPSTSVSLTTAVQDYVCPQLQSHALCTCCFQ 398
QY 51 ---GPAGLGABEPAAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLR----- 99
Db 399 PMPDRAEREQDPRAVQ-----QCAVC---LQPFCHLYWGCTRTGCGYGLA 442
QY 100 -----SLGAVVFSRVTVNNVLEAPFLVGIEGSLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
Db 443 PFCELNLGDKLDGVLNNNSYEDILKNYLAT-RGLTWKNML-----TES 486
QY 154 ALAALRGHFCLS-----SDKMVCYLLKTKAIVNASEMDIQNVPLSE 194
Db 487 LVALORGVELLSYRVGTGTVLCYCCGLRSFRELTYQYRQNPASE 532

RESULT 13
AAB93182
ID AAB93182 standard; protein; 623 AA.
AC AAB93182;
XX
XX 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12128.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX
XX EF1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX

PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 12128; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95993
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 623 AA;

Query Match 7.4%; Score 89; DB 4; Length 623;
Best Local Similarity 23.0%; Pred. No. 2.4;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPPRGDFC-----GTERAIDQA-----SFTTSMEDTQVVGSSPL----- 50
Db 393 YRRQAQPPH---CPAPEGEPGAPQALGDAPSTSVSLTTAVQDYVCPQLQSHALCTCCFQ 449
QY 51 ---GPAGLGABEPAAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLR----- 99
Db 450 PMPDRAEREQDPRAVQ-----QCAVC---LQPFCHLYWGCTRTGCGYGLA 493
QY 100 -----SLGAVVFSRVTVNNVLEAPFLVGIEGSLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
Db 494 PFCELNLGDKLDGVLNNNSYEDILKNYLAT-RGLTWKNML-----TES 537
QY 154 ALAALRGHFCLS-----SDKMVCYLLKTKAIVNASEMDIQNVPLSE 194
Db 538 LVALORGVELLSYRVGTGTVLCYCCGLRSFRELTYQYRQNPASE 583

RESULT 14
AAB97233
ID AAB97233 standard; protein; 623 AA.
XX
XX AAB97233;
XX
XX 27-JUN-2002 (first entry)

XX DE Novel human protein SEQ ID NO: 501.
XX KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
XX KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX KW neuroprotective; antiparkinsonian; protein therapy; EST;
XX KW expressed sequence tag.
XX OS Homo sapiens.
XX PN WO200222660-A2.
XX PD 21-MAR-2002.
XX PF 10-SEP-2001; 2001WO-US026015.
XX PR 11-SEP-2000; 2000US-00659671.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX DR WPI; 2002-292408/33.
XX DR N-PSDB; ABN32419.
XX PT An isolated polynucleotide for treating diseases associated with its
XX PT encoded polypeptide such as cancer and multiple sclerosis.
XX PS Example 2; SEQ ID NO 501; 509pp; English.
XX CC The present invention provides the protein and coding sequences of 444
XX CC novel human proteins. These were isolated from expressed sequences tags
XX CC (ESTs). They can be used to stimulate cell growth, to regulate
XX CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX CC e.g. in burn treatment, to regulate the immune system e.g. to treat
XX CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX CC rheumatoid arthritis, and to treat nervous system disorders e.g.
XX CC Parkinson's disease. The present sequence is a protein of the invention
XX CC
XX CC Sequence 623 AA;
XX CC
XX CC Query Match 7.4%; Score 89; DB 5; Length 623;
XX CC Best Local Similarity 23.0%; Pred. No. 2.4;
XX CC Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
XX CC
XX CC 8 HRSRCATPPRGDFC-----GGERALDQA-----SFTTSMEDTQVVGSSPL----- 50
XX CC 393 YRQAAQPPH---CPAPEGEPAPQALGDAPPTSVSLTTAVQDYVCPLOGSHALCTCCFQ 449
XX CC 51 ---GPAGLGAEBEPAAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLR----- 99
XX CC 450 PMPDRAEREQDPRVAPQ-----QCAVC---LQPFCHLYWGCTRTGCGYGCLA 493
XX CC 100 -----SLGAVVFSRVTVNNVLEAPFLVIGIEGSLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
XX CC 494 PFCELNLGDKCLDGLVNLNNSYESDILKNYLAT-RGLTWKNML-----TES 537
XX CC 154 ALAALRGHCFLS-----SDKWCYLLKTRAIYNASEMDIQNYPLSE 194
XX CC 538 LVALQRGVLLSDYRVGTDTVLVCCGLRSFRELTYOQNIPASE 583
XX CC
XX CC RESULT 15
XX CC AAB93168
XX CC ID AAB93168 standard; protein; 652 AA.
XX CC
XX CC AAB93168;
XX CC
XX CC 26-JUN-2001 (first entry)
XX CC

DE Human protein sequence SEQ ID NO:12100.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX DR WPI; 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesising 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesising polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX CC
XX CC Sequence 652 AA;
XX CC
XX CC Query Match 7.4%; Score 89; DB 4; Length 652;
XX CC Best Local Similarity 23.0%; Pred. No. 2.5;
XX CC Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
XX CC
XX CC 8 HRSRCATPPRGDFC-----GGERALDQA-----SFTTSMEDTQVVGSSPL----- 50
XX CC 422 YRQAAQPPH---CPAPEGEPAPQALGDAPPTSVSLTTAVQDYVCPLOGSHALCTCCFQ 478
XX CC 51 ---GPAGLGAEBEPAAGPOLPSWLQPERCAVFOCAOCHAVLADSVHLAWDLR----- 99
XX CC 479 PMPDRAEREQDPRVAPQ-----QCAVC---LQPFCHLYWGCTRTGCGYGCLA 522
XX CC 100 -----SLGAVVFSRVTVNNVLEAPFLVIGIEGSLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
XX CC 523 PFCELNLGDKCLDGLVNLNNSYESDILKNYLAT-RGLTWKNML-----TES 566

Qy 154 ALAALRGHCLS-----SDXWCVLLKTKAIVNASEMDIONVPLSE 194
Db 567 LVALQRGVFLLSDYRVGTDTVLCYCCGLRSFRELTYQYRONIPASE 612

Search completed: May 27, 2005, 08:49:07
Job time : 163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 08:43:58 ; Search time 43 Seconds
(without alignments)
397.550 Million cell updates/sec

Title: US-09-942-052A-728
Perfect score: 1198
Sequence: 1 MAQPLRHSRCATPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	7.2	158	4	US-09-134-000C-6596
2	83	6.9	457	6	Sequence 6596, Ap Patent No. 5268463
3	83	6.9	457	6	Patent No. 5268463
4	81.5	6.8	456	6	Patent No. 5432081
5	81.5	6.8	456	6	Patent No. 5432081
6	81.5	6.8	457	2	US-08-882-704A-6
7	81.5	6.8	457	4	US-09-151-957-6
8	80.5	6.7	1711	2	US-08-342-930-2
9	78.5	6.6	325	2	US-09-018-576-3
10	78.5	6.6	325	2	US-09-018-576-12
11	78.5	6.6	325	3	US-09-248-137-3
12	78.5	6.6	325	3	US-09-248-137-12
13	78	6.5	190	4	US-09-252-991A-19049
14	77	6.4	607	3	US-09-537-682-1
15	77	6.4	719	2	US-08-520-933-3
16	77	6.4	719	4	US-09-285-040-3
17	77	6.4	738	6	5262177-2
18	77	6.4	738	6	5262177-2
19	77	6.4	776	4	US-09-949-016-10829
20	76.5	6.4	282	4	US-09-845-713A-2
21	76	6.3	351	4	US-09-902-540-13214
22	75.5	6.3	310	1	US-08-129-456A-36
23	75.5	6.3	310	2	US-08-705-868-3
24	75.5	6.3	310	3	US-09-123-615-3
25	75.5	6.3	310	3	US-08-360-821B-35
26	75.5	6.3	310	4	US-09-851-026-35
27	75.5	6.3	1449	3	US-08-840-062-6

28	75.5	6.3	1723	4	US-09-194-612A-31	Sequence 31, Appl
29	75.5	6.3	4545	2	US-08-804-227C-14	Sequence 14, Appl
30	75.5	6.3	4550	2	US-08-804-227C-8	Sequence 8, Appl
31	75.5	6.3	4550	2	US-08-804-198-2	Sequence 2, Appl
32	75	6.3	78	4	US-09-270-767-43496	Sequence 43496, A
33	75	6.3	458	4	US-09-800-170-3	Sequence 3, Appl
34	75	6.3	717	6	5262177-5	Patent No. 5262177
35	75	6.3	717	6	5262177-5	Patent No. 5262177
36	75	6.3	784	4	US-09-724-586A-1	Sequence 1, Appl
37	75	6.3	784	4	US-09-600-823-1	Sequence 1, Appl
38	75	6.3	784	4	US-09-724-666-1	Sequence 1, Appl
39	75	6.3	1088	4	US-09-920-804-2	Sequence 2, Appl
40	75	6.3	1088	4	US-09-920-804-4	Sequence 4, Appl
41	75	6.3	1129	4	US-09-734-674-2	Sequence 2, Appl
42	75	6.3	2227	3	US-08-475-886-4	Sequence 4, Appl
43	75	6.3	2227	4	US-09-653-499-4	Sequence 4, Appl
44	75	6.3	2227	4	US-10-135-988-4	Sequence 4, Appl
45	74.5	6.2	211	4	US-09-949-016-11409	Sequence 11409, A

ALIGNMENTS

RESULT 1
US-09-134-000C-6596
; Sequence 6596, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6596
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6596

Query Match 7.2%; Score 86; DB 4; Length 158;
Best Local Similarity 26.5%; Pred. No. 0.16;
Matches 40; Conservative 27; Mismatches 54; Indels 30; Gaps 9;
QY 87 LADSVHLAWDLSESLGAVV---FSRVTVNNVLEA--PELVGIEGSLKG-----STYN 133
Db 17 LTYIYILASEQPIUSTGSLTGLSKPTGNLKNKAPFFINDFEGDPGKFSYNYLTKLN 76
QY 134 LFLFCGCGIPVGFPHLYSTHAALAAAL---RGHFCLSSDKMVCYLLKTKAIVNASEMDIQNV 190
Db 77 YQVCSNCSLPSDNQNTETPSAIALGVIKRNHF-----KNTC---TTKLYILSALNGPENK 128
QY 191 PLSE-----KIAELKEKIVLTHNRLKSLMKIL 217
Db 129 PISKTVTRKLSLSYK-DLYYQEMK-LLKII 157

RESULT 2
5268463-7
; Patent No. 5268463
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: PLAAAT PROMOTER a-GLUCURONIDASE GENE
; CONSTRUCT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,976
; FILING DATE: 08-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 119,102

; FILING DATE: 10-NOV-1987
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; SEQ ID NO:7:
; LENGTH: 457
5268463-7

Query Match 6.9%; Score 83; DB 6; Length 457;
Best Local Similarity 25.2%; Pred. No. 1.8;
Matches 37; Conservative 30; Mismatches 54; Indels 26; Gaps 8;

QY 95 WDLRSGLGAVVFSRVTVNNV-----VLEA-----PFLVGIEGSLKGYSTYNLL-FCGSCG 141
Db WSLPVALVALAIAASIQGGVTMTVMWALEADTVEYGEYLTGVR--IEGLTYSLSFSTRKCG 373
QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 374 QAIGGS--SIPAFILGSLGYIANQVOTPEVINGIRTSIALVPCGFMLLAFVIWIFYPLTD 431
QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221
Db 432 K--KFKETVVEIDNRKKVQQQLISDIT 456

RESULT 3

5268463-7
; Patent No. 5268463
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: PLAANT PROMOTER a-GLUCURONIDASE GENE
; CONSTRUCT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,976
; FILING DATE: 08-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 119,102
; FILING DATE: 10-NOV-1987
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; SEQ ID NO:7:
; LENGTH: 457
5268463-7

Query Match 6.9%; Score 83; DB 6; Length 457;
Best Local Similarity 25.2%; Pred. No. 1.8;
Matches 37; Conservative 30; Mismatches 54; Indels 26; Gaps 8;

QY 95 WDLRSGLGAVVFSRVTVNNV-----VLEA-----PFLVGIEGSLKGYSTYNLL-FCGSCG 141
Db WSLPVALVALAIAASIQGGVTMTVMWALEADTVEYGEYLTGVR--IEGLTYSLSFSTRKCG 373
QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 374 QAIGGS--SIPAFILGSLGYIANQVOTPEVINGIRTSIALVPCGFMLLAFVIWIFYPLTD 431
QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221
Db 432 K--KFKETVVEIDNRKKVQQQLISDIT 456

RESULT 4

5432081-7
; Patent No. 5432081
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: HOST CELLS TRANSFORMED WITH THE E. COLI
; GLUCORONIDE PERMEASE GENE
; CURRENT APPLICATION DATA:
; NUMBER OF SEQUENCES: 10
; APPLICATION NUMBER: US/08/138,546
; FILING DATE: 15-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 447,976
; FILING DATE: 08-DEC-1989

; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; APLICATION NUMBER: 119,102
; FILING DATE: 10-NOV-1987
; SEQ ID NO:7:
; LENGTH: 456
5432081-7

Query Match 6.8%; Score 81.5; DB 6; Length 456;
Best Local Similarity 25.2%; Pred. No. 2.6;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSGLGAVVFSRVTVNNV-----VLEA-----PFLVGIEGSLKGYSTYNLL-FCGSCG 141
Db WSLPVALVALAIAASIQGGVTMTVMWALEADTVEYGEYLTGVR--IEGLTYSLSFSTRKCG 373
QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 374 QAIG--GSIPAFILGSLGYIANQVOTPEVINGIRTSIALVPCGFMLLAFVIWIFYPLTD 430
QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221
Db 431 K--KFKETVVEIDNRKKVQQQLISDIT 455

RESULT 5

5432081-7
; Patent No. 5432081
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: HOST CELLS TRANSFORMED WITH THE E. COLI
; GLUCORONIDE PERMEASE GENE
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,546
; FILING DATE: 15-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 447,976
; FILING DATE: 08-DEC-1989
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; APPLICATION NUMBER: 119,102
; FILING DATE: 10-NOV-1987
; SEQ ID NO:7:
; LENGTH: 456
5432081-7

Query Match 6.8%; Score 81.5; DB 6; Length 456;
Best Local Similarity 25.2%; Pred. No. 2.6;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSGLGAVVFSRVTVNNV-----VLEA-----PFLVGIEGSLKGYSTYNLL-FCGSCG 141
Db WSLPVALVALAIAASIQGGVTMTVMWALEADTVEYGEYLTGVR--IEGLTYSLSFSTRKCG 373
QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 374 QAIG--GSIPAFILGSLGYIANQVOTPEVINGIRTSIALVPCGFMLLAFVIWIFYPLTD 430
QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221
Db 431 K--KFKETVVEIDNRKKVQQQLISDIT 455

RESULT 6

US-08-882-704A-6
; Sequence 6, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; APPLICANT: Wilson, Katherine J.
; APPLICANT: Leader, Michael
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SEED and BERRY LLP
;; STREET: 6300 Columbia Center, 701 Fifth Avenue
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: USA
;; ZIP: 98104-7092
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/882,704A
;; FILING DATE: 25-JUN-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5879906tenburg Ph.D., Carol
;; REGISTRATION NUMBER: 39,317
;; REFERENCE/DOCKET NUMBER: 190106.404
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 457 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-882-704A-6

Query Match 6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 2.6;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFVSRTNNV-----VLEA-----PFLVGIEGSLKGTNNLL-FCGSCG 141
DB 317 WSLPVALVALAIASIGQVTTVMWALEADTVVEYGEYLTGVR--IEGLTYSLSFSTRKCG 374
QY 142 IPVGFLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNAASEMDIQNV-----PLSE 194
DB 375 QAIG--GSIPAFILGSGVIANQVTPPEVINGIRTSIALVPCGFALLAFVIWFFPLTD 431
QY 195 KTAELKEKIVLTHNRLKSLMKILSEVT 221
DB 432 K--KPKKEIVVEIDNRKKVQQLISDIT 456

RESULT 7
US-09-151-957-6
; Sequence 6, Application US/09151957
; Patent No. 6429292
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; Leader, Michael
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,957
; FILING DATE: 11-Sep-1998

;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/882,704
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 6429292tenburg Ph.D., Carol
;; REGISTRATION NUMBER: 39,317
;; REFERENCE/DOCKET NUMBER: 190106.404
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 457 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-151-957-6

Query Match 6.8%; Score 81.5; DB 4; Length 457;
Best Local Similarity 25.2%; Pred. No. 2.6;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFVSRTNNV-----VLEA-----PFLVGIEGSLKGTNNLL-FCGSCG 141
DB 317 WSLPVALVALAIASIGQVTTVMWALEADTVVEYGEYLTGVR--IEGLTYSLSFSTRKCG 374
QY 142 IPVGFLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNAASEMDIQNV-----PLSE 194
DB 375 QAIG--GSIPAFILGSGVIANQVTPPEVINGIRTSIALVPCGFALLAFVIWFFPLTD 431
QY 195 KTAELKEKIVLTHNRLKSLMKILSEVT 221
DB 432 K--KPKKEIVVEIDNRKKVQQLISDIT 456

RESULT 8
US-08-342-930-2
; Sequence 2, Application US/08342930
; Patent No. 5821084
; GENERAL INFORMATION:
; APPLICANT: OLMSTED, ELIZABETH A.
; APPLICANT: MAURO, LAURA J.
; APPLICANT: DAVIS, ALAN R.
; APPLICANT: DIXON, JACK E.
; TITLE OF INVENTION: OSTEOBLAST-TESTICULAR PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/342,930
; FILING DATE: 21-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOSKI, ANTOINETTE F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 20344-20975.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141

QY 225 S 225
Db 237 S 237

RESULT 11
US-09-248-137-3
; Sequence 3, Application US/09248137
; Patent No. 6030788
; GENERAL INFORMATION:
; APPLICANT: Gerhold, David L.
; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, RY60-30
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/248,137
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,576
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19885Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-248-137-3

Query Match 6.6%; Score 78.5; DB 3; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.4;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
QY 22 GTERAIDQASFTTSMEDTQVVKSSPLGPAGLG-----AEEPAAG 63
Db 43 GPNQALREIKALQEME-DNQYVQLKAVFPHGCGFVLAFEFMLSLDAEVVTHAQRLPAQ 101
QY 64 POLPSWLOPERCAVFCQACQACHAVLADSVHLAWDLRSGLGVFSRVYNNV-----LEAPP 119
Db 102 AQVKSYLQMLLKGV---AFCHA-----NNIVHRLKPAUL 133
QY 120 LVGIEGSLKSTYNNLFCGSCGIPVGFHLYSTHAALALRGHFCCLSD-----167
Db 134 LISASQQLKIADFLGARVFS---PDGSRLY--THQ--VATRSVGCIMGELLINGSPLPPGKN 187
QY 168 --KMVCYLLKTKAI VNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQ 224
Db 188 DIEQLCYVLRILGTPNPQVMPPELTLPDYNKIS-FKEQVMP-----LEEVLDPV 236
QY 225 S 225
Db 237 S 237

RESULT 13
US-09-252-991A-19049
; Sequence 19049, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

RESULT 12
US-09-248-137-12
; Sequence 12, Application US/09248137
; Patent No. 6030788
; GENERAL INFORMATION:
; APPLICANT: Gerhold, David L.
; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, RY60-30
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/248,137
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,576
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19885Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-248-137-12

Query Match 6.6%; Score 78.5; DB 3; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.4;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
QY 22 GTERAIDQASFTTSMEDTQVVKSSPLGPAGLG-----AEEPAAG 63
Db 43 GPNQALREIKALQEME-DNQYVQLKAVFPHGCGFVLAFEFMLSLDAEVVTHAQRLPAQ 101
QY 64 POLPSWLOPERCAVFCQACQACHAVLADSVHLAWDLRSGLGVFSRVYNNV-----LEAPP 119
Db 102 AQVKSYLQMLLKGV---AFCHA-----NNIVHRLKPAUL 133
QY 120 LVGIEGSLKSTYNNLFCGSCGIPVGFHLYSTHAALALRGHFCCLSD-----167
Db 134 LISASQQLKIADFLGARVFS---PDGSRLY--THQ--VATRSVGCIMGELLINGSPLPPGKN 187
QY 168 --KMVCYLLKTKAI VNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQ 224
Db 188 DIEQLCYVLRILGTPNPQVMPPELTLPDYNKIS-FKEQVMP-----LEEVLDPV 236
QY 225 S 225
Db 237 S 237

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19049
LENGTH: 190
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19049

Query Match 6.5%; Score 78; DB 4; Length 190;
Best Local Similarity 32.9%; Pred. No. 1.7;
Matches 24; Conservative 6; Mismatches 29; Indels 14; Gaps 3;
QY 4 QPLRHRSRCATPRGDFCGTERRAIDQASFTTSMEDWDTQVVGSSPLGPAAGLGAEEPAAG 63
Db 94 RPECHTRCVRPAAGYPG---RAGDRAS-----VARSTGELPAAPGRPRGHC- 139
QY 64 PQLPSWLQPERCA 76
Db 140 PAPPWPWRRCRA 152

RESULT 14
US-09-537-682-1
Sequence 1, Application US/09537682
Patent No. 6303357
GENERAL INFORMATION:
APPLICANT: TAKEUCHI, Kenichi
APPLICANT: KOIDE, Yoshinao
APPLICANT: NAKANISHI, Yuji
APPLICANT: SUZUKI, Satoru
TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE, RECOMBINANT
TITLE OF INVENTION: DNA, AND METHOD FOR PRODUCING MODIFIED
TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE
FILE REFERENCE: A20-121814C/KI
CURRENT APPLICATION NUMBER: US/09/537,682
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 607
TYPE: PRT
ORGANISM: Enterococcus faecium No. 6303357 7044
US-09-537-682-1

Query Match 6.4%; Score 77; DB 3; Length 607;
Best Local Similarity 23.2%; Pred. No. 13;
Matches 49; Conservative 30; Mismatches 76; Indels 56; Gaps 10;
QY 19 DFCGGTERRAIDQASFTTSM-----EWDQVVGSSPLGP-----AGLGAEEPAAGPQ 65
Db 51 DPAEGT-----SSRSTKLHGIRYLKTFDVEV--ADTVGERAVVQGIAPHIPKPDPM 102
QY 66 LPSWLQPERCAVQCAQCHAVLADSLRSILGAVVFSRVTN-----NVVLEAPF 119
Db 103 LLPVDEGATTNNMF-----SVKVAMDLYDKLANVTGKYENYTLTPEVLEREPF 154
QY 120 LVGIEGSLKGSYNLLFCSCGIPVGF-----HLVSTHAALALRGHFLCSSDKMVCYLL 174
Db 155 LK--KEGLKA-----GYLDFRNNDARLVIDNIKKAEDGAYLVSKWKAAGFLY 202
QY 175 KTKAIVNASEMDIQNVPLSEKIAELKEKIVL 205
Db 203 EGDQIVGVKARDL-----LTDEVIEIKSLVI 229

RESULT 15
US-08-520-933-3
Sequence 3, Application US/08520933
Patent No. 5981194
GENERAL INFORMATION:
APPLICANT: Jefferies, Wilfred A.
APPLICANT: McGeer, Patrick L.
APPLICANT: Rothenberger, Sylvia
APPLICANT: Food, Michael R.
APPLICANT: Yamada, Tatsuo
APPLICANT: Kennard, Malcolm
TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/520,933
FILING DATE: August 31, 1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Shona S. McDiarmid
REGISTRATION NUMBER: 38,798
REFERENCE/DOCKET NUMBER: 7685-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-364-7311
TELEFAX: 416-361-1398
TELEX: 06-23115
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-520-933-3

Query Match 6.4%; Score 77; DB 2; Length 719;
Best Local Similarity 23.7%; Pred. No. 17;
Matches 59; Conservative 25; Mismatches 87; Indels 78; Gaps 15;
QY 17 RGDGCGTERRAIDQASFTTSMEW-----DTQVVGSSPLGPAGLGAEEPAAGP 64
Db 171 RGDSSG--EGVCDKSPLEYYDYGAFRCLAGAGDVAFVRHSTVL-----ENTDOK 220
QY 65 QLPSSLQF-----ERCAVFOCAQCH--AVLADSVHLAWDLRSILGAVVFSRV 109
Db 221 TLPSWGQALLSQDFELLCDGSRADVTETWROCHLARVPAHAVVVRADTD---GGLIF-RL 276
QY 110 TNNVLEAPFLVGIGGS-----LKGSTY---NLLFCGSG--IPVGFHLYST---HAALAA 157
Db 277 LN-----EQQLFSHSGSSQFMSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHYLHA 332
QY 158 LRGHFCLSSDKMVCYL-----LKTKAIVNASEMDIQNVPLSEKIAELKEKIVLTHNRLKSL 213
Db 333 MKGLLC--DPNRLPPYLRWCVLSTPEIQKQGM-----AVAFRRRLRKP 375
QY 214 MKILSEVTP 222
Db 376 IQCVSAKSP 384

Search completed: May 27, 2005, 08:53:45

Job time : 45 secs

BEST AVAILABLE COPY

This Page Blank (uspto)